FORMA Y MODALIDAD

Una introducción al concepto de consecuencia lógica

Mario Gómez Torrente

eudeba

eL

enciclopedia lógica

Esta colección incluye distintos trabajos que exponen los principales problemas filosóficos presentes en la lógica contemporánea. La idea es que cada volumen presente el contenido de un concepto central de la mencionada disciplina. En todos los casos, no sólo se intenta describir el estado actual de las discusiones alrededor de ese concepto, sino además ofrecer un punto de vista original.

Forma y modalidad es una excelente introducción al concepto de *consecuencia lógica*, en especial, a su enfoque tarskiano y a la problemática filosófica que se ha generado en los últimos tiempos alrededor de esa definición. En este libro se describen los principales intentos de analizar las propiedades formales y modales del mencionado concepto, poniendo énfasis particular, en la idea tarskiana según la cual la idea de *consecuencia lógica* tiene que ser reconstruida a partir de la noción de *preservación de verdad por toda interpretación*. Mario Gómez Torrente defiende muy convincentemente la tesis de que el enfoque tarskiano ofrece una buena reconstrucción de los aspectos modales y formales del concepto intuitivo de *consecuencia lógica*. Para ello explora, de una manera muy clara y precisa, las principales objeciones en contra de esta idea sembrando agudas dudas acerca de la posibilidad de encontrar buenos motivos para descartarla. El libro logra lo que se propone: ser sencillo y al mismo tiempo profundo.

Eduardo Alejandro Barrio
Director de la colección

eudeba
www.eudeba.com.ar

ISBN 950-23-1087-X el0050

Forma y modalidad

Una introducción al concepto de consecuencia lógica

Mario Gómez Torrente

Eudeba

Eudeba
Universidad de Buenos Aires

1ª edición: diciembre de 2000

© 2000
Editorial Universitaria de Buenos Aires
Sociedad de Economía Mixta
Av. Rivadavia 1571/73 (1033) - Ciudad de Buenos Aires
Tel.: 4383-8025 / Fax: 4383-2202
www.eudeba.com.ar

Diseño de tapa: *Juan Cruz Gonella*
Diagramación: *Félix C. Lucas*
Corrección general: Eudeba

ISBN 950-23-1087-X
Impreso en Argentina.
Hecho el depósito que establece la ley 11.723

No se permite la reproducción total o parcial de este libro, ni su almacenamiento en un sistema informático, ni su transmisión en cualquier forma o por cualquier medio, electrónico, mecánico, fotocopia u otros métodos, sin el permiso previo del editor.

Enciclopedia Lógica

Los argumentos están presentes en nuestras conversaciones, debates, discusiones y hasta en algunas de nuestras ficciones y propagandas publicitarias. Desde sus inicios, la lógica se ocupa de diferenciar los buenos de los malos argumentos, intentando ofrecernos criterios para reconstruir la idea de *convencimiento racional*. Se supone que nuestros buenos argumentos no nos convencen por la fuerza: nuestra aceptación de los enunciados que ellos apoyan, técnicamente, nuestra aceptación de la conclusión del argumento, es el resultado de un proceso en el cual nos hemos persuadido de que otros enunciados, técnicamente las premisas del argumento, son suficientes para producir dicha aceptación. En breve, nuestro convencimiento es el resultado de que la conclusión se desprende de las premisas. Es clave, entonces, entender en qué consiste este desprendimiento. Tradicionalmente, los buenos argumentos han sido llamados *correctos*. En un argumento correcto, la aceptación de la conclusión es una *consecuencia* necesaria de la aceptación de las premisas. Intuitivamente, esta idea ha sido abordada desde dos planos:

En el plano semántico, quiere decir que un enunciado es una consecuencia lógica de un conjunto de premisas si y solamente si es imposible que todas las premisas resulten verdaderas y la conclusión no lo sea. O lo que es lo mismo, un enunciado será una consecuencia lógica de un conjunto de enunciados si y sólo si es necesario que si todas las premisas son verdaderas la conclusión también lo sea. Sólo en este caso, la verdad de las premisas justifica la verdad de la conclusión.

Esta misma idea, en el plano sintáctico ha sido tratada revalorizando las propiedades formales o estructurales de los argumentos. Así, un enunciado es una consecuencia lógica de un conjunto de premisas si y sólo si hay una secuencia finita de enunciados tales cada uno de los integrantes de esa secuencia es o bien una premisa o bien se sigue de enunciados que le preceden en la secuencia en función de la aplicación de una regla de inferencia primitiva. En este caso, se dice que la conclusión del argumento se deriva de las premisas de ese argumento.

Ambos planos son vistos en este volumen de la colección *Enciclopedia lógica* por Mario Gómez Torrente como marcando las propiedades modales y formales del concepto de *consecuencia lógica*. El tema principal del libro el enfoque tarskiano y a la problemática filosófica que se ha generado en los últimos tiempos alrededor de esa definición. Mario Gómez Torrente pone énfasis particular, en la idea tarskiana según la cual la idea de *consecuencia lógica* tiene que ser reconstruida a partir de la noción de *preservación de verdad por toda interpretación*. A partir de esa definición formal que toma fundamentalmente un camino semántico, defiende muy convincentemente la tesis de que el enfoque tarskiano ofrece una buena reconstrucción de los aspectos modales y formales del concepto intuitivo de *consecuencia lógica*. Para ello explora, de una manera muy clara y precisa, las principales objeciones en contra de esta idea sembrando agudas dudas acerca de la posibilidad de encontrar buenos motivos para descartarla. El libro logra lo que se propone: ser sencillo y al mismo tiempo profundo.

Este es la quinta publicación de la colección. Anteriormente, hemos publicado *Lógica informal* de Juan Comesaña que ofrece otros criterios, que los analizados en el presente volumen, para evaluar argumentos. La idea principal que se defiende, de manera clara y precisa, es que el mero hecho de que un argumento tenga un parecido de familia con un tipo de razonamiento tradicionalmente clasificado como falaz, no implica que ese argumento no pueda ser legítimamente usado en una discusión racional. Hemos publicado *La verdad desestructurada,* escrito por mí, en donde presento las ideas tarskianas acerca de la verdad y argumento en contra del tradicional enfoque correspondentista. También hemos editado *Concepciones de la referencia* donde Eleonora Orlando ofrece una respuesta a fondo al desafío que en nuestros días presentan las corrientes anti-representacionalistas. La idea principal que se defiende, de una manera original y muy convincente, es que la relación de *referencia* debe

explicarse en términos mixtos, esto es, haciendo uso tanto de aspectos descriptivos como de relaciones causales. Finalmente, hemos publicado *Conjuntos e infinitos* en el cual Carolina Sartorio analiza la noción intuitiva de *conjunto* y se presentan algunas de las principales teorías actuales acerca de esa noción. El libro contiene un desarrollo preciso de las paradojas de *conjuntos demasiado grandes* y de las asombrosas ideas de Cantor al respecto. Quiero agradecer a las autoridades de EUDEBA por la confianza y el estimulante apoyo recibido.

Eduardo Alejandro Barrio
Director de la Colección

Índice

Prefacio 11

I. El concepto de consecuencia lógica 13

II. Formalidad, modalidad y teorías de la consecuencia lógica 21

III. Consecuencia lógica y derivabilidad 27

IV. La teoría de Tarski en "Sobre el concepto de consecuencia lógica" 35

V. Teoría, definición y corrección 43

VI. La definición tarskiana usual de consecuencia lógica 51

VII. Consecuencia lógica y generalidad 57

VIII. Nociones generalistas de modalidad y lenguajes de segundo orden 67

IX. Consecuencia lógica, aprioridad, analiticidad 77

X. El concepto de constante lógica y sus dificultades 85

XI. La definición de constante lógica de Tarski 93

XII. Un examen del concepto tarskiano de constante lógica 101

Apéndice. Las semicomillas 107

Referencias bibliográficas 109

Prefacio

La monografía que el lector tiene en sus manos es una introducción a los temas más fundamentales de la reflexión filosófica reciente sobre el concepto de consecuencia lógica y, especialmente, a la definición tarskiana de consecuencia lógica y a la discusión filosófica sobre esta definición. El texto es una versión ampliada de los materiales en borrador sobre el concepto de consecuencia lógica preparados para un seminario con el título "Verdad y Consecuencia Lógica" impartido en la Facultad de Filosofía y Letras de la Universidad de Buenos Aires, en el segundo semestre de 1996. Mi intención principal al escribir esta versión ampliada ha sido crear un texto de referencia al que los lectores pudiesen acudir en busca de una exposición de los temas, ideas y conceptos más básicos relacionados con la noción de consecuencia lógica; por ello, el nivel de la exposición quiere ser elemental, y los únicos conocimientos que se presuponen son los que da un curso básico de lógica de primer orden que se acompañe de una presentación de las nociones elementales de la teoría de conjuntos. Espero, sin embargo, que la monografía pueda ser de algún provecho para otras personas con un interés más especializado en el concepto de consecuencia lógica y en la noción, estrechamente relacionada con éste, de constante lógica. Con este público en mente he incluido discusiones sobre algunos temas algo más avanzados que los meramente introductorios. Estas discusiones aparecen en los capítulos marcados con un asterisco (*). Su lectura no es imprescindible para la comprensión de

los otros capítulos. (Pero recomiendo al no especialista que, si es posible, lea los capítulos marcados con un asterisco después de haber leído los otros capítulos.) Al público especialista me permitiría recomendarle también la lectura de mis trabajos mencionados en la sección de referencias bibliográficas, donde muchas cuestiones relacionadas con los temas de esta monografía, pero no discutidas en ella, se tratan a un nivel comparativamente avanzado (y donde la gran mayoría de las cuestiones elementales tratadas con cierto detalle en esta monografía no se tratan en absoluto, o se tratan con la negligencia deliberada típica de los trabajos especializados).

Agradezco a Sandra Lazzer que me propusiera para impartir el seminario mencionado más arriba y que me estimulara a preparar los materiales que finalmente se convertirían en esta monografía. También agradezco a Alberto Moretti haber leído una versión anterior y haber propuesto algunos cambios. A Eduardo Barrio le doy las gracias por su interés en publicar la monografía y por sus sugerencias sobre cómo ampliar los materiales originales.

I. EL CONCEPTO DE CONSECUENCIA LÓGICA

El concepto de *consecuencia* se usa con mucha frecuencia en un gran número de ámbitos. Estos usos suelen ser de gran importancia, por ejemplo en contextos políticos, jurídicos y científicos. Un ciudadano puede reprochar al representante elegido por su circunscripción que es una *consecuencia* de las afirmaciones de éste en la campaña electoral que, de ser elegido, votaría a favor de cierta ley (a pesar de que no lo ha hecho); el político, naturalmente, puede replicar que no *se sigue* tal cosa de sus afirmaciones preelectorales. En un caso legal disputado, el fiscal puede sostener la tesis de que *se sigue* de la ley vigente y de los precedentes pertinentes que el acusado ha de recibir cierto castigo; el abogado defensor puede argumentar que, por el contrario, no es una *consecuencia* de la ley y de los precedentes pertinentes que su cliente haya siquiera de ser castigado. Un sociólogo puede mantener que es una *consecuencia* de la teoría sociológica que ha defendido siempre (quizá conjugada con una serie de premisas empíricas adicionales) que cierto fenómeno social o político recién ocurrido había de ocurrir; otro sociólogo puede poner esa afirmación en duda.

Estos ejemplos muestran no sólo la importancia del concepto de consecuencia (o de sus usos), sino también la deseabilidad de una comprensión lo más adecuada posible de ese concepto. Una comprensión más adecuada que la comprensión común quizá no permita la resolución de disputas como las de los ejemplos, pero es sin duda condición necesaria para una posible resolución, y es deseable en sí misma.

Desde la antigüedad, los filósofos han procurado iluminar nuestra comprensión del concepto de consecuencia. También desde la antigüedad, los logros más importantes de estas reflexiones filosóficas se han basado en la suposición (a veces tácita y no argumentada, a veces explícita y defendida) de que hay un concepto especial de consecuencia con propiedades particulares, concepto que algunos han llamado de *consecuencia lógica*. Entre las propiedades atribuidas, a veces sólo tácitamente, al concepto de consecuencia lógica es importante destacar dos.

La primera de estas propiedades del concepto de consecuencia lógica es que la *relación* que determina entre conjuntos de oraciones (premisas) y oraciones que son consecuencias lógicas de esas premisas (conclusiones), la relación de consecuencia lógica, es una relación *formal*. Que esta relación es formal quiere decir lo siguiente: si una oración A es una consecuencia lógica de un conjunto de oraciones B, C, D, etc., entonces cualquier argumento que tenga *la misma forma* que el argumento con premisas B, C, D, etc., y conclusión A será también un argumento en que la conclusión es una consecuencia lógica de las premisas. A estas alturas de nuestra exposición no es posible decir mucho más sobre la noción de forma (y de identidad de forma), pero un ejemplo servirá para ilustrar la idea intuitiva en que se basa.

Consideremos el argumento con premisas 'Si la teoría cuántica es verdadera entonces el hombre es un ser libre' y 'La teoría cuántica es verdadera', y con conclusión 'El hombre es un ser libre', que podemos representar, como se hace habitualmente, de la siguiente manera:

Si la teoría cuántica es verdadera entonces el hombre es un ser libre
La teoría cuántica es verdadera

El hombre es un ser libre.

En este argumento, la conclusión parece una consecuencia de las premisas. Si aceptamos que además es una consecuencia lógica de ellas y que otros argumentos tienen la misma forma, estos argumentos también serán ejemplos de consecuencia lógica; por ejemplo, si aceptamos que el argumento

Si Juan se comió el helado de la nevera entonces Doña Pura se disgustará
Juan se comió el helado de la nevera

Doña Pura se disgustará

tiene la misma forma que el primer argumento, entonces la conclusión también se seguirá lógicamente de las premisas en este caso.

No parece que el concepto general de consecuencia tenga la propiedad de determinar una relación formal, a diferencia del concepto especial de consecuencia lógica. Por ejemplo, podría sostenerse que de la premisa 'El hombre es libre' se sigue (aunque no lógicamente) la conclusión 'El hombre tiene responsabilidad moral'. Sin embargo el argumento

El avestruz es comestible

El avestruz tiene sensibilidad estética

parecería tener la misma forma sin que su conclusión se siga (lógicamente o no) de su premisa. Naturalmente, la idea de que el concepto general de consecuencia no determina una relación formal depende, para su justificación, de una clarificación previa de la noción de forma que no tenemos a mano ahora. En el capítulo X daremos una mínima clarificación de la noción de forma, clarificación que depende de varias ideas que no hemos expuesto aún, y que tiene como resultado que los dos argumentos mencionados en este párrafo tienen la misma forma. (Desde otros puntos de vista sobre la noción de forma podría mantenerse que esos argumentos tienen distinta forma, aunque esos puntos de vista son infrecuentes.) Pero antes de llegar a esa clarificación haremos varias suposiciones típicas sobre la forma de varios argumentos; todas ellas resultarán familiares para el lector acostumbrado a la formalización de argumentos en lógica.

La segunda propiedad importante atribuida al concepto de consecuencia lógica es que la relación que determina entre premisas y conclusiones es una relación *modal* bastante estricta. Es modal en cuanto que es una relación de implicación necesaria, donde la conclusión se sigue *necesariamente* de las premisas. Aristóteles, que utilizó por primera vez la idea de que la relación de

consecuencia lógica es formal, fue también el primero en afirmar que es una relación modal: "El silogismo es un discurso en el que, una vez enunciadas ciertas cosas, se sigue necesariamente algo distinto de lo ya enunciado del hecho de que aquellas cosas se den" (*Primeros Analíticos*, 24b: 18-20).

Esta relación modal es bastante estricta en el sentido de que, cuando se da entre aserciones, esto es una condición suficiente para que se den otros tipos de implicación necesaria entre esas aserciones, a pesar de que no es condición necesaria. Por ejemplo, en filosofía es bastante común distinguir entre implicación por necesidad lógica e implicación por necesidad metafísica, y, a su vez, entre estos tipos de implicación y la implicación por necesidad física y otras. Un ejemplo de implicación por necesidad lógica sería la implicación de las premisas 'Si Juan se comió el helado de la nevera entonces Doña Pura se disgustará' y 'Juan se comió el helado de la nevera' a la conclusión 'Doña Pura se disgustará'; siendo lógicamente necesaria, es también metafísicamente necesaria, físicamente necesaria, etc. Pero las relaciones inversas no se dan. Podemos conceder que el enunciado 'Juan es un hombre' implica por necesidad metafísica el enunciado 'Juan es animal' (y también que este enunciado es una consecuencia –aunque no sea lógica– de aquel), pero esta implicación –se concede también– no se da por necesidad lógica. Podemos admitir que el enunciado 'Juan es un hombre' implica por necesidad física el enunciado 'Juan no vuela' (y también que este enunciado se sigue de aquel), pero esta implicación no se da por necesidad lógica (y seguramente tampoco metafísica).

¿Coinciden la relación de implicación por necesidad lógica y la relación de consecuencia lógica? ¿Son la misma relación? El punto de vista usual sobre esta cuestión es que son relaciones distintas. Se suele aceptar, por ejemplo, que ciertas implicaciones que se dan en virtud del significado de ciertos predicados son implicaciones por necesidad lógica, a pesar de que no son casos de consecuencia lógica. Usando una ilustración muy socorrida, se diría que 'Doña Pura es soltera' implica por necesidad lógica 'Doña Pura es una mujer no casada'; sin embargo, se diría también que 'Doña Pura es una mujer no casada' no es una consecuencia lógica de 'Doña Pura es soltera'. La justificación de esta afirmación podría apelar al hecho de que otros argumentos con la misma forma no son casos de consecuencia lógica; por ejemplo, el argumento

Doña Pura es pescadera

Doña Pura es una motocicleta no engrasada

no sería un caso de consecuencia lógica (ni de consecuencia a secas), a pesar de tener la misma forma que el otro argumento. De nuevo el ejemplo se basa en ciertas suposiciones acerca de la forma de los argumentos.

Los ejemplos de los dos últimos párrafos sirven también para hacer plausible la tesis adicional de que el concepto general de consecuencia y el concepto especial de consecuencia lógica son diferentes, al determinar relaciones diferentes. Parece perfectamente aceptable decir que de que Doña Pura es soltera *se sigue* que Doña Pura es una mujer no casada, si bien acabamos de decir que no se sigue lógicamente. También parece aceptable decir que de que Juan es un hombre *se sigue* que Juan no vuela. Pero esta implicación no es necesaria en el mismo sentido en que lo es la implicación de las premisas 'Si Juan se comió el helado de la nevera entonces Doña Pura se disgustará' y 'Juan se comió el helado de la nevera' a la conclusión 'Doña Pura se disgustará'. Esta implicación, suponemos, es un caso de consecuencia lógica, pero la anterior, al no ser una implicación por necesidad lógica, no es un caso de consecuencia lógica, aunque sea un caso de consecuencia.

Un punto de vista alternativo sobre estos ejemplos consiste en negar que los supuestos casos de consecuencia no lógica sean de hecho casos de consecuencia. Desde este punto de vista, el concepto de consecuencia y el concepto de consecuencia lógica coincidirían (al menos extensionalmente). También desde este punto de vista, el hecho aparente de que 'Doña Pura es una mujer no casada' se sigue de 'Doña Pura es soltera' y 'Juan no vuela' se sigue de 'Juan es un hombre' se explica proponiendo que lo que en realidad ocurre es que 'Doña Pura es una mujer no casada' se sigue *lógicamente* de 'Doña Pura es soltera' *junto con* otras premisas verdaderas –por ejemplo, una definición de 'soltera' por un filólogo–; y que 'Juan no vuela' se sigue *lógicamente* de 'Juan es un hombre' *junto con* varias premisas verdaderas –por ejemplo, leyes y aserciones observacionales de la física y la biología–. En esta concepción, argumentos como

Juan es un hombre

Juan no vuela

no son ejemplos de consecuencia, si bien son *entimemas*, argumentos donde se han omitido varias premisas verdaderas y que se convierten en argumentos lógicamente correctos (es decir, en ejemplos de consecuencia lógica) una vez que esas premisas son restituidas.

Es posible también sostener un punto de vista estrechamente relacionado con este último. Es posible sostener que argumentos como el de 'Juan es un hombre' a 'Juan no vuela' *son* ejemplos de consecuencia, si bien son *reducibles* a o *reconstruibles* como entimemas cuyas premisas adicionales son verdaderas. Este punto de vista no está exento de problemas. Uno de ellos es que toda verdad es la conclusión de un entimema con premisas verdaderas, y sin embargo podría pensarse que no todo argumento con una conclusión verdadera es un caso de consecuencia.

Problemas aparte, el hecho de que puntos de vista como los descritos en los dos párrafos precedentes hayan sido adoptados con bastante frecuencia entre los filósofos y el hecho de que tengan una cierta plausibilidad, ilustran la importancia de la noción especial de consecuencia lógica: desde ambos puntos de vista la noción general de consecuencia se reduce en último término a la noción de consecuencia lógica (y a otras nociones que se suponen menos problemáticas, como la de verdad). Por ésta y otras razones, la investigación filosófica sobre el concepto de consecuencia se ha concentrado sobre el concepto especial de consecuencia lógica. La construcción de una teoría o descripción iluminadora y adecuada de este concepto y de su extensión –la relación de consecuencia lógica– ha sido el objetivo fundamental de la investigación lógica y filosófica sobre la noción de consecuencia.

La mayor parte de esta monografía está dedicada al estudio de una teoría sobre la relación de consecuencia lógica, la teoría formulada por el lógico Alfred Tarski en 1936. Esta teoría ha sido ampliamente aceptada por lógicos y filósofos y sólo recientemente algunos filósofos y lógicos han ofrecido argumentos que concluyen que es inadecuada. (Sin embargo, nadie ha puesto en duda que sea muy iluminadora.) Dada la complejidad de la mayoría de esos argumentos, en esta monografía apenas podemos discutir y criticar algunos.

En Gómez Torrente (1998/9), que recomiendo al lector interesado, defiendo detenidamente la tesis de que esos argumentos, aunque motivados por preocupaciones legítimas, no son correctos, y sugiero que no se ha mostrado que la teoría de Tarski sea inadecuada, al menos por el momento.

En el capítulo II describiremos de una forma somera y resumida las principales aproximaciones históricas a la relación de consecuencia lógica como intentos de aprehender sus cualidades de formalidad y modalidad. Este repaso histórico nos servirá, luego, para introducir en el capítulo III algunas de las motivaciones específicas de Tarski para proponer su teoría, teoría que enunciaremos y explicaremos en los capítulos IV, V y VI. En los capítulos V, VI, VII, VIII y IX estudiaremos con cierto detalle la cuestión de si la teoría de Tarski es correcta. Finalmente, en los capítulos X, XI y XII exploraremos algunas interrelaciones entre la noción de forma y la cuestión de la corrección de la teoría de Tarski.

II. Formalidad, modalidad y teorías de la consecuencia lógica

Una explicación o teoría adecuada de la relación de consecuencia lógica ha de acomodar o respetar las dos propiedades básicas de esa relación que mencionamos en el capítulo I: el ser *formal* y el ser *modal*. El hecho mismo de que hayamos aislado estas propiedades podría sugerirnos una primera teoría de la relación de consecuencia lógica. Podríamos proponer como nuestra teoría la tesis de que la relación de consecuencia lógica se da entre un conjunto de premisas P_1, P_2, P_3, etc., y una conclusión C exactamente cuando P_1, P_2, P_3, etc., *implican por necesidad lógica* C y, además, todo argumento *con la misma forma* que el argumento

P_1
P_2
P_3
etc.

C

es también una implicación por necesidad lógica.

No hay nada obviamente erróneo en proponer una tesis semejante, pero desde un punto de vista filosófico se podrían poner en duda su capacidad explicativa y su utilidad. Una dificultad preliminar tendría que ver con el

hecho de que la tesis a secas, tal como la hemos enunciado, no se basa en una teoría que nos permita identificar la forma de argumentos particulares (y, por tanto, ser capaces siquiera de preguntarnos si un argumento es un ejemplo de consecuencia lógica según la teoría). Esta dificultad podría paliarse si el ofrecedor de la tesis propusiese también una serie de tesis suplementarias sobre cómo desvelar las formas de argumentos particulares. La objeción principal a la tesis, sin embargo, es que no explica o intenta delimitar la noción de consecuencia lógica en términos de nociones mejor comprendidas y más claras que ella. En particular, es muy cuestionable que nuestra comprensión del concepto de implicación por necesidad lógica sea superior a nuestra comprensión del concepto de consecuencia lógica. Naturalmente, ambos son conceptos altamente intuitivos que no se hallan respaldados por teorías o conjuntos de tesis bien establecidas más allá de las simples convicciones ordinarias sobre ellos. Pero ésta no sería una crítica pertinente a la tesis si el concepto de implicación por necesidad lógica fuese más claro o mejor comprendido que el de consecuencia lógica. El problema es que nuestra comprensión de ambos conceptos intuitivos parece *igualmente* pobre; y así, cuando consideramos la cuestión de si una cierta oración es implicada por necesidad lógica por un cierto conjunto de oraciones, no se nos sugiere ninguna forma de evaluarla que sea sustancialmente diferente a la que intuitivamente aplicaríamos si nos hiciéramos la pregunta de si el argumento correspondiente es un caso de consecuencia lógica.

Aristóteles, el fundador de la lógica como disciplina, ofreció una primera aproximación fructífera a la combinación de las nociones de formalidad y modalidad lógica en el concepto de consecuencia lógica. Aristóteles aisló un buen número de formas de argumentos, ofreciendo de este modo una caracterización precisa de la noción de forma en cuanto restringida a esa clase de argumentos. Además, observó que los argumentos de unas cuantas formas básicas de las aisladas por él eran silogismos o argumentos donde la conclusión se seguía necesariamente de las premisas. Y, por último, observó también que la implicación por necesidad que se da en los argumentos de otras cuantas formas podía justificarse a partir de la de las formas básicas por aplicaciones sucesivas de reglas de transformación que intuitivamente producían argumentos por implicación necesaria cuando se aplicaban a argumentos por implicación necesaria. De este modo, Aristóteles caracterizó de una forma muy precisa un conjunto bastante amplio de argumentos lógicamente correctos:

los argumentos de las formas identificadas por él como básicas más los obtenibles por sucesivas aplicaciones de sus reglas de transformación a partir de argumentos de las formas básicas. (Aristóteles quizá creyó que este conjunto de argumentos era *el* conjunto de los argumentos correctos, si bien esta interpretación de sus opiniones es discutible.)

Los filósofos estoicos y megáricos aislaron muchas nuevas formas argumentales (que hoy se estudian en lógica proposicional o de enunciados), identificaron unas cuantas formas básicas de implicación por necesidad lógica y al parecer propusieron un cálculo de transformaciones inspirado en la misma idea que el de Aristóteles. Tanto Aristóteles como los estoicos y megáricos ofrecieron teorías muy precisas e iluminadoras de la relación de consecuencia lógica. En primer lugar, hicieron precisa la noción de forma para los tipos de argumentos de que se ocupaban. En segundo lugar, explicaron o elucidaron la noción de necesidad lógica usando sistemas axiomáticos, apelando a la idea de que una gran cantidad de formas argumentales de implicación necesaria (quizá todas) podrían generarse empleando un número escaso de intuiciones incontrovertibles acerca de la necesidad de unas pocas implicaciones básicas y la transmisibilidad de esa necesidad por unas pocas reglas de transformación.

Gottlob Frege, lógico de importancia comparable a la de Aristóteles y fundador de la lógica moderna a finales del siglo XIX, se aproximó a la relación de consecuencia lógica de un modo análogo, en esencia, al aristotélico, estoico y megárico: Frege identificó muy importantes nuevas formas argumentales y ofreció un cálculo axiomático para acotar un conjunto de formas lógicamente correctas. Pero muchas de sus contribuciones fueron sin embargo revolucionarias, con respecto a la lógica anterior, y de tremenda profundidad. En primer lugar, Frege inventó un lenguaje simbólico (o una serie de lenguajes), diseñado especialmente para la reproducción de argumentos matemáticos, en el cual siempre es enteramente claro cuál es la forma de un argumento y si dos argumentos tienen la misma forma o no. Frege conjeturó además que, cuando menos, todo argumento matemático podría formalizarse por medio de un argumento en su lenguaje simbólico. El lenguaje que inventó Frege era lo que hoy llamaríamos un lenguaje cuantificacional de orden superior. Este lenguaje contenía como fragmento lo que hoy llamaríamos un lenguaje cuantificacional de primer orden, un lenguaje como los que se estudian hoy en los cursos básicos de lógica. (En estos lenguajes la

identificación de las importantes formas argumentales aisladas por primera vez por Frege se hace especialmente perspicua.)

En segundo lugar, Frege incrementó de manera formidable el rigor de la presentación axiomática en un cálculo lógico, hasta el punto de que puede considerársele el creador de la idea de sistema formal: él fue el primero en ofrecer un sistema formal como los que estudiamos en los cursos básicos de lógica. En el sistema de Frege, se acepta que las oraciones de unas pocas formas oracionales básicas (en su lenguaje formalizado), los axiomas, son lógicamente necesarias, y se dan dos únicas reglas de transformación o inferencia que, cuando se aplican a oraciones, generan oraciones que son implicadas por necesidad lógica por aquellas.

Una vez construido un sistema así, es posible ofrecer una caracterización muy precisa de un conjunto de argumentos lógicamente correctos en el lenguaje formalizado del sistema: podemos proponer que la relación de consecuencia lógica se da entre un conjunto de premisas P_1, P_2, P_3, etc., y una conclusión C (del lenguaje del sistema) exactamente cuando existe una serie de aplicaciones de las reglas de inferencia que, partiendo de P_1, P_2, P_3, etc., y posiblemente también de oraciones de las formas axiomáticas básicas, acaba en C. Cuando una serie tal existe se dice que C es *derivable* o *demostrable* o *deducible* en el sistema formal a partir de P_1, P_2, P_3, etc.

Esta tesis respeta el requisito de que la relación de consecuencia lógica sea formal y modal. En primer lugar, la relación de derivabilidad en un sistema como el de Frege o los de los manuales de lógica es formal, pues si una conclusión C es derivable de un conjunto de premisas P_1, P_2, P_3, etc., en uno de esos sistemas, y si un argumento con premisas P_1', P_2', P_3', etc., y conclusión C' tiene la misma forma que el primero, entonces C' es también derivable del conjunto P_1', P_2', P_3', etc. (usando una serie de derivación isomorfa a la existente para el primer argumento).

En segundo lugar, la relación de derivabilidad en un sistema como el de Frege o los de los cursos de lógica es una relación modal exactamente en el sentido deseado. Las oraciones que se toman como axiomas son todas ellas lógicamente necesarias, implicadas por necesidad lógica por cualquier oración o conjunto de oraciones; y las reglas de transformación o inferencia siempre generan oraciones implicadas por necesidad lógica por las oraciones a las que se aplican. (Veremos una ligera salvedad a estas afirmaciones en el capítulo

IX.) El resultado es que partiendo de un conjunto de premisas y aplicando las reglas de transformación a las premisas, a axiomas o a oraciones generadas previamente de esta guisa, las únicas oraciones que será posible derivar serán oraciones que las premisas implican por necesidad lógica.

Observemos además que nuestra comprensión de la relación de derivabilidad en un sistema como el de Frege es, sin duda, mejor y más clara que nuestra comprensión de los conceptos de consecuencia lógica y de implicación por necesidad lógica. La pregunta de si una conclusión es derivable en el sistema a partir de un conjunto de premisas es una pregunta esencialmente más precisa y mejor delimitada que la pregunta llana de si esa conclusión es implicada por necesidad lógica por las premisas o de si es una consecuencia lógica de ellas. El acercamiento a la relación de consecuencia lógica en términos de la de derivabilidad en ciertos sistemas goza, por tanto, de un gran atractivo metodológico y explicativo. En gran parte por estos motivos, este acercamiento proporcionaría la concepción dominante de la relación de consecuencia lógica entre los lógicos durante largo tiempo, especialmente desde la publicación de la importante obra de Russell y Whitehead *Principia Mathematica* (1910-1913), que en muchos sentidos es una refinación y extensión del enfoque de Frege.

En 1936, Tarski, en su conocido artículo "Sobre el concepto de consecuencia lógica" (Tarski, 1936), propuso una teoría sobre la relación de consecuencia lógica apoyada en una base explicativa enteramente diferente a la de las teorías basadas en la noción de derivabilidad. La teoría de Tarski no incluye innovaciones con respecto a la manera de hacer perspicua la forma de los argumentos; Tarski ofrece su teoría con miras a su aplicación a los *mismos* lenguajes formales creados por Frege, y aceptando la noción de forma para oraciones y argumentos de esos lenguajes implícita en Frege; estos lenguajes son, en esencia, los lenguajes que hoy llamamos cuantificacionales clásicos de órdenes primero y superiores. La innovación de Tarski tiene que ver con el modo en que su teoría incorpora el componente modal de la relación de consecuencia lógica; en la teoría de Tarski, este componente no se intenta recoger apelando a las propiedades de una cierta relación de derivabilidad, sino a un aparato de ideas completamente diferente.

El principal motivo de Tarski para ofrecer su teoría era su creencia de que, en ciertos lenguajes formales de los inventados por Frege, ningún cálculo de axiomas y reglas de transformación con las características del fregeano o el de

Principia Mathematica bastaría para derivar todas las consecuencias lógicas de todos los posibles conjuntos de premisas. Digámoslo de manera ligeramente más precisa: Tarski observó que, para cierto lenguaje formal cuantificacional de orden superior L, y para cualquier cálculo *Cal* con las propiedades benignas de un cálculo como el de Frege (es decir, con axiomas especificables de forma efectiva y reglas de inferencia también especificables y aplicables de forma efectiva, y que no permitan la derivación de contradicciones), existirá un conjunto de premisas P_1, P_2, P_3, etc., y una conclusión C en L tales que C es una consecuencia lógica intuitiva de P_1, P_2, P_3, etc., y sin embargo C no será derivable de P_1, P_2, P_3, etc., por medio del cálculo *Cal*.

No podemos entrar con mucho más detalle en la observación de Tarski y en su discusión sin presuponer ciertos conocimientos no elementales de lógica. Al lector con los conocimientos requeridos está destinado el siguiente capítulo.

III. Consecuencia lógica y derivabilidad*

En "Sobre el concepto de consecuencia lógica", Tarski dice que los lógicos habían pensado hasta tiempos recientes que habían conseguido definir un concepto preciso que coincidía en extensión con el concepto intuitivo de consecuencia lógica. Tarski menciona el extraordinario desarrollo de la lógica en las décadas anteriores, que había mostrado "cómo presentar las disciplinas matemáticas en la forma de teorías deductivas formalizadas" (Tarski, 1936: 409). Es decir, teorías como las ideadas por Frege, donde las consecuencias se extraen de axiomas y teoremas por medio de reglas de inferencia puramente sintácticas y finitarias, como por ejemplo las reglas de *modus ponens* o de particularización universal que se estudian en los cursos básicos de lógica. "Siempre que una oración se sigue de otras, puede obtenerse a partir de ellas –eso se creía– por medio de transformaciones prescritas por las reglas" (Tarski, 1936: 410). Según Tarski, esta creencia de los lógicos se justificaba por "el hecho de que habían tenido éxito en la tarea de reproducir en la forma de pruebas formalizadas todos los razonamientos exactos que habían sido llevados a cabo en la matemática" (Tarski, 1936: 410).

Pero Tarski dice que la creencia de los lógicos estaba equivocada. Hay algunos casos en que una cierta oración se sigue en el sentido intuitivo de un

*Este capítulo trata temas algo más avanzados que los introductorios; puede omitirse en una primera lectura.

conjunto de oraciones pero no puede ser derivada a partir de esas oraciones usando los axiomas y reglas aceptados. Estos casos los proporcionan algunas teorías ω-incompletas, teorías en las que para algún predicado P las oraciones

$P(0)$,
$P(1)$,
$P(2)$,

y, en general, todas las oraciones de la forma $P(n)$ (donde en lugar de 'n' vaya un numeral que denote a un número natural) pueden derivarse, pero la oración universal

Para todo número natural n, P(n)

no puede derivarse usando los axiomas y reglas de inferencia aceptados.

> Sin embargo, intuitivamente parece cierto que la oración universal ['Para todo número natural n, $P(n)$'] se sigue en el sentido usual de la totalidad de oraciones particulares [$P(0)$, $P(1)$, $P(2)$,...]. Supuesto que todas estas oraciones sean verdaderas, la oración ['Para todo número natural n, $P(n)$'] también ha de ser verdadera (Tarski, 1936: 411).

Tarski considera la posibilidad de añadir una regla ω a las reglas aceptadas de inferencia, esto es, una regla que nos permita derivar una oración universal de la forma de 'Para todo número natural n, $P(n)$' a partir del conjunto de oraciones $P(0)$, $P(1)$, $P(2)$,... Sin embargo, dice también que una regla semejante diferiría en aspectos esenciales de las reglas usuales: no es una regla finitaria, mientras que todas las reglas aceptadas en los sistemas deductivos habituales son finitarias. Luego Tarski señala que podría enunciarse una cierta versión finitaria de la regla para algunas teorías en las que se pueda construir suficiente aritmética. Sea $B(P)$ una oración aritmética de la teoría que codifique (por medio de una numeración de Gödel) el enunciado de que todos los casos numerales particulares de un predicado P pueden ser derivados en la teoría por medio de un conjunto dado de axiomas y reglas; y sea $C(P)$ la generalización universal de ese predicado (relativizada a los números

naturales). Entonces la regla que permite derivar $C(P)$ de $B(P)$ es una regla bastante complicada de enunciar, pero es una regla finitaria para el lenguaje de la teoría en la que se puede construir suficiente aritmética, y una regla que podemos añadir al conjunto de reglas de inferencia dado. (Obsérvese que el hecho de que todos los casos numerales particulares de un predicado P sean derivables por medio de las reglas dadas no implica por sí solo que la oración $B(P)$ sea derivable usando las reglas dadas.)

Pero Tarski quita inmediatamente importancia a la sugerencia de extender el sistema de reglas dado con una regla ω formalizada y finitaria. Señala que, en vista de los resultados de incompleción de Gödel, por muchas reglas o axiomas finitarios que añadamos a la teoría y a extensiones finitarias de la teoría, las teorías resultantes seguirán siendo incompletas, y de hecho ω-incompletas. Estas observaciones bastan, según Tarski, para mostrar que "si queremos obtener el concepto de consecuencia adecuado, que esté cercano en lo esencial al concepto común, debemos recurrir a métodos muy diferentes y aplicar un aparato conceptual muy diferente al definirlo" (Tarski, 1936: 413). Estos métodos y este aparato conceptual son los métodos desarrollados por el propio Tarski, con el objeto de dar sus conocidas definiciones de un concepto preciso de verdad y de otros conceptos semánticos. La aplicación de estos métodos en la definición de un concepto preciso de consecuencia lógica se describe detenidamente en los capítulos IV y VI de esta monografía.

El ejemplo de las teorías ω-incompletas es el único que propone Tarski para motivar la necesidad de una teoría del concepto de consecuencia lógica formulada en términos de nociones distintas a la de derivabilidad. Este ejemplo ha causado alguna perplejidad en la bibliografía sobre estos temas. La razón de esta perplejidad es que, en las formalizaciones usuales de la aritmética de primer orden, una oración de la forma

Para todo número natural n, P(n)

no es una consecuencia (en el sentido definido por Tarski, que se expone luego en esta monografía) del conjunto de oraciones de las formas $P(0)$, $P(1)$, $P(2)$,... (Esto es así porque, como quedará claro luego, el predicado 'es un número natural' y los numerales pueden recibir interpretaciones distintas de la usual, y no en todas ellas es el caso que alguna de las premisas sea falsa o la conclusión

verdadera.) Puede añadirse más perplejidad si consideramos el hecho de que, para los lenguajes de primer orden, el teorema de compleción de Gödel nos garantiza que la relación de consecuencia definida por Tarski y la relación normal de derivabilidad coinciden en extensión. John Etchemendy ha expresado esta perplejidad del siguiente modo:

> "los ejemplos de Tarski tienen que ver con la relación de consecuencia para lenguajes de primer orden, donde la relación [definida por Tarski] coincide con la relación definida sintácticamente. ¿Cómo puede juzgarse a una teoría [como la de Tarski] superior a la caracterización sintáctica usual si las dos son, de hecho, extensionalmente equivalentes?" (Etchemendy, 1990: 85).

La perplejidad se resuelve observando que, cuando dio su ejemplo de las teorías ω-incompletas, Tarski no estaba afirmando la corrección lógica de una versión de la regla ω para el lenguaje de la aritmética de primer orden. Estaba afirmando la validez lógica de una versión de esta regla, sí, pero una versión de esta regla para el lenguaje de la teoría simple de los tipos, un lenguaje con cuantificación sobre variables de todos los tipos finitos. Este lenguaje contiene únicamente signos primitivos cuya naturaleza lógica es aventurado poner en duda (conectivas, cuantificadores sobre variables de todos los tipos finitos y variables de esos mismos tipos). En él, es posible definir los conceptos aritméticos al estilo logicista de Frege o Russell y formular proposiciones de la aritmética elemental en términos de esos conceptos. Los axiomas aritméticos pueden derivarse como verdades expresadas de la manera logicista, a partir de axiomas característicos de la teoría de tipos –tales como axiomas de comprensión para cada tipo, de extensionalidad para cada tipo y de infinitud.

Es una consecuencia del teorema de incompleción de Gödel que, para cualquier cálculo efectivo *Cal* para un lenguaje tal con axiomas y reglas consistentes, habrá argumentos con infinitas premisas $P(0)$, $P(1)$, $P(2)$,... (donde P es un predicado de los números naturales) y conclusión 'Para todo número natural n, $P(n)$' tales que la conclusión no es derivable de las premisas por medio de *Cal*. Sin embargo, estos argumentos son intuitivamente correctos, por lo menos en cuanto que, dada la definición logicista de los números, todos los argumentos de la forma

P(0)
P(1)
P(2)
etc.

Para todo número natural n, P(n)

serán casos de implicación por necesidad lógica. Obsérvese que, dada la definición logicista usual de los números naturales como ciertas clases de clases biyectables de individuos, no hay ninguna interpretación posible del predicado P en que todas las premisas de uno de esos argumentos sean verdaderas y su conclusión falsa. (Esto será así independientemente de cuál sea la cardinalidad que se le supone al universo de los individuos, aunque, naturalmente, sólo si este universo es infinito será infinito también el conjunto de los "números" definidos lógicamente. Si el universo de los individuos es finito, el conjunto de los "números" definidos lógicamente será finito, pero seguirá siendo la clase denotada por el predicado lógicamente definido 'es un número natural'.) Intuitivamente, las propiedades de los primitivos empleados en la definición logicista de las nociones aritméticas han de ser responsables del hecho de que haya consecuencia lógica en estos casos. Sin embargo, ningún cálculo puede recoger adecuadamente esas propiedades y plasmarlas por medio de la noción de derivabilidad. De ahí el interés de Tarski por crear un nuevo tipo de teoría.

No cabe duda de que Tarski consideraba los primitivos habituales de la teoría simple de tipos como expresiones lógicas, y, por tanto, a sus propiedades inferenciales como objeto de estudio de la lógica. Un texto de la época donde esto queda especialmente claro se halla en el manual de lógica que Tarski publicó en 1937:

> "resulta que la noción misma de número e igualmente todos los otros conceptos aritméticos son definibles dentro del ámbito de la lógica. De hecho, es fácil establecer el significado de los símbolos que designan números naturales, tales como '0', '1', '2', etc. El número 1, por ejemplo, puede definirse como el número de elementos de una clase que consta exactamente de un elemento. (...) Y tampoco es difícil definir el concepto general de número natural: un número natural es el número cardinal de una clase finita. Además, nos

hallamos en la posición de definir todas las operaciones sobre números naturales, y de extender el concepto de número con la introducción de fracciones, números negativos y números irracionales, sin ir en ningún momento más allá de los límites de la lógica. Además, es posible demostrar todos los teoremas de la aritmética sobre la base de las leyes de la lógica solamente (...). Esta construcción es muy abstracta, no puede popularizarse fácilmente y no cabe dentro de una presentación elemental de la aritmética (...). Pero el mero hecho de que haya sido posible desarrollar toda la aritmética, incluyendo las disciplinas erigidas sobre ella –el álgebra, el análisis, etc.– como una parte de la lógica pura, constituye uno de los más grandes logros de las investigaciones lógicas recientes" (Tarski, 1937: 50-51).

La postura de Tarski, que convierte a los primitivos del lenguaje de la teoría de tipos en expresiones cuyas propiedades inferenciales son objeto de estudio para la lógica, no es tan frecuente hoy en día como lo era en la época en que escribió Tarski. Sin embargo, la postura de que al menos los cuantificadores de los tipos más bajos son expresiones lógicas sigue siendo una postura defendida por muchos lógicos y filósofos (véanse, por ejemplo, Boolos, 1975, 1984 y Shapiro, 1991). No hay un consenso que haya hecho que se acepten universalmente los argumentos de quienes niegan que los cuantificadores de los lenguajes de orden superior sean expresiones lógicas o que formas argumentales tradicionalmente aceptadas como correctas en la lógica de orden superior sean casos de implicación por necesidad lógica. Algunos de estos argumentos se discuten y rechazan en el capítulo IX de este trabajo. Para una discusión más detallada de argumentos de este tipo, véase Gómez Torrente (1998/9).

Concluyamos este capítulo examinando brevemente una objeción que podría plantearse a la concepción informal que Tarski tenía de la regla ω. Esta objeción podría plantearse así: si las inferencias permitidas por una versión de la regla ω son correctas, entonces deberían ser lógicamente correctas consideradas bajo cualquier formalización de la aritmética, se trate de una formalización a la manera logicista o de una formalización en primer orden donde los numerales y el predicado 'es un número natural' no se toman como símbolos definidos; por tanto –concluiría la objeción–, la definición de consecuencia lógica de Tarski ha de ser incorrecta, pues no declara lógicamente correctas a aquellas inferencias en una formalización de primer orden, donde se entiende que aquellas

expresiones son reinterpretables de manera diferente en interpretaciones distintas a la interpretación aritmética normal.

Esta objeción no está mejor fundada que la tesis falsa de que un argumento correcto ha de ser lógicamente correcto bajo cualquiera de sus formalizaciones. Una observación familiar en los cursos introductorios de lógica es que algunos argumentos correctos son lógicamente correctos bajo algunas de sus formalizaciones, pero no bajo otras, menos refinadas. Así,

Todos los griegos son hombres
Todos los hombres son mortales

Todos los griegos son mortales

es un argumento correcto, pero sólo algunas de sus formalizaciones son lógicamente correctas; por ejemplo, el argumento

p
q

---

r

de un lenguaje proposicional no es lógicamente correcto, pero el argumento

$\forall x(Gx \rightarrow Hx)$
$\forall x(Hx \rightarrow Mx)$

$\forall x(Gx \rightarrow Mx)$

de un lenguaje de primer orden es lógicamente correcto. Esta es una buena manera de ver que ni nosotros ni Tarski tenemos ninguna obligación de aceptar la tesis de que si una versión de la regla ω es lógicamente correcta entonces todas sus versiones –y, en particular, sus versiones en primer orden– son lógicamente correctas también. Sólo de algunas versiones de la regla ω –y, en particular, de algunas de sus versiones en la teoría simple de los tipos– es razonable pensar que son lógicamente correctas; y ésta fue, sin duda, la postura de Tarski.

IV. La teoría de Tarski en "Sobre el concepto de consecuencia lógica"

La mejor forma de apreciar la naturaleza de la teoría tarskiana de la consecuencia lógica es desarrollarla paso a paso, viendo los resultados que da para un lenguaje formalizado particular de estructura simple. El lenguaje que escogemos será un lenguaje de primer orden para un fragmento muy sencillo de la aritmética elemental. Este lenguaje, *LAr*, tiene signos primitivos para el cuantificador universal de primer orden ('$\forall$'), el condicional ('$\rightarrow$'), la negación ('$\neg$'), paréntesis ('(', ')') (el cuantificador existencial de primer orden, '$\exists$', y las otras conectivas veritativo-funcionales, como '$\wedge$' y '$\vee$', se pueden definir de formas acostumbradas usando estos símbolos), igualdad ('='), la letra 'x', un acento subíndice ('.') para generar infinitas variables por posposición a 'x' (para abreviar, nos referiremos a la variable en que la letra 'x' va seguida de *n* acentos subíndices como 'x_n'), la constante individual '0', el predicado monádico 'N', y el predicado diádico 'M'. Ejemplos de enunciados de este lenguaje y sus significados son 'N0' ("0 es un número natural"), '$\forall x(Nx \rightarrow \neg Mx0)$' ("para todo número natural x, x no es menor que 0"), '$\forall x(Nx \rightarrow \exists x.(Nx. \wedge Mxx.))$' ("para todo número natural x existe un número natural x. tal que x es menor que x.").

Para un lenguaje como *LAr*, Tarski haría en 1936 una cierta suposición acerca del significado de los cuantificadores. En las oraciones aritméticas de *LAr*, los cuantificadores en una oración han de estar siempre relativizados al predicado 'N'; es decir, un cuantificador universal siempre significa "para todo objeto singular cualquiera..." y un cuantificador existencial siempre significa

"existe un objeto...", no "para todo número natural" y "existe un número natural...". Así, las oraciones '$\forall x(Nx \rightarrow \neg Mx0)$' y '$\forall x(\neg Mx0)$' significan cosas distintas. La primera es aritmética, con cuantificadores relativizados, y es verdadera. La segunda no es puramente aritmética, es acerca de *objetos singulares cualesquiera* (y es falsa –pues, por ejemplo, -1 es menor que 0–). Tarski no adopta la práctica, común hoy en día, de restringir implícitamente el recorrido de las variables de un lenguaje formalizado al conjunto de objetos de la interpretación deseada de ese lenguaje (véase Gómez Torrente, 1996). Para conseguir ese resultado, los lenguajes formalizados han de contar con predicados que, como 'N', permitan restringir apropiadamente el recorrido de las variables a través del proceso de relativización de los cuantificadores.

En las consideraciones preliminares a la propuesta de su teoría, Tarski dice que cuando una oración X de un lenguaje formal (por ejemplo, *LAr*) es una consecuencia lógica de un conjunto K de oraciones de ese lenguaje, el argumento con premisa K y conclusión X tiene la siguiente propiedad, que Tarski llama 'condición (F)':

> "(F) Si, en las oraciones del conjunto K y en la oración X, las constantes –aparte de las constantes puramente lógicas– son sustituidas por cualesquiera otras constantes (con los mismos signos siempre sustituidos por los mismos signos), y si llamamos 'K'" al conjunto así obtenido a partir de K, y 'X'" a la oración obtenida a partir de X, entonces la oración X' debe ser verdadera dado solamente que todas las oraciones de K' sean verdaderas" (Tarski, 1936: 415).

Aclaremos el sentido de la condición (F) por medio de un ejemplo. Consideremos un lenguaje LAr^+, que es como *LAr* pero además tiene otra constante individual, '2', otro predicado monádico, 'Ps', y otro predicado diádico, 'Pd', cuyas interpretaciones deseadas son el número 2, la propiedad "ser un entero positivo" y la propiedad "ser predecesor inmediato de", respectivamente. Sea K el siguiente conjunto de oraciones de LAr^+: {'$\forall x(Nx \rightarrow \neg Mx0)$', 'N0'} (estas oraciones son verdaderas); y sea X la oración '$\neg M00$'. El argumento con premisa K y conclusión X es, intuitivamente, lógicamente correcto (además, al tener premisas verdaderas, su conclusión ha de ser verdadera, y en efecto podemos comprobar que lo es); por tanto, según Tarski, ha de cumplir la condición

(F). Esto quiere decir que *cualquier* argumento obtenido a partir de él por sustitución uniforme de constantes no lógicas por constantes no lógicas ha de ser un argumento donde *no es el caso que las premisas sean verdaderas y la conclusión falsa*. Supongamos que las constantes no lógicas de LAr^+ son '0', 'N', 'M', '2', 'Ps' y 'Pd'. Sustituyamos '0' por '2', 'N' por 'Ps' y 'M' por 'Pd' en el argumento con premisa K y conclusión X y llamemos al conjunto de premisas y la oración resultantes 'K'" y 'X'". Es decir, K' es {'$\forall x(Psx \rightarrow \neg Pdx2)$', 'Ps2'} y X' es '$\neg Pd22$'. En virtud de la condición (F), el argumento de K' a X' ha de ser un argumento donde no es el caso que las premisas sean verdaderas y la conclusión falsa; y en efecto es así: la primera premisa es falsa y la segunda premisa y la conclusión son verdaderas.

Que todo argumento lógicamente correcto satisface la condición (F) se sigue de la propiedad de formalidad de los argumentos lógicamente correctos que enunciamos en el capítulo I, junto con cierta suposición sobre la noción de forma de un argumento u oración. La propiedad de formalidad dice que si un argumento es lógicamente correcto, entonces todo argumento de la misma forma lo es también. La suposición sobre la noción de forma es una explicación de esta noción en términos de la noción de constante lógica, explicación que nosotros aceptaremos también en esta monografía; la suposición es la siguiente: dos argumentos (no necesariamente distintos) tienen la misma forma cuando son obtenibles el uno a partir del otro sustituyendo, de manera uniforme, constantes no lógicas por constantes no lógicas. Así, pues, si un argumento con premisa K y conclusión X es lógicamente correcto, la propiedad de formalidad y la suposición recién mencionada implican que un argumento con premisa K' y conclusión X' es lógicamente correcto y, por tanto, al ser un caso de implicación por necesidad lógica, que no es el caso que sus premisas sean verdaderas y su conclusión falsa.

Tarski se pregunta si es posible ofrecer la condición (F) como *definición* de la relación de consecuencia lógica, es decir, si es aceptable proponer que la condición (F) es no sólo una condición necesaria sino, también, una condición suficiente para que un argumento sea un caso de consecuencia lógica. Su respuesta es que no es posible. La razón de esto es que la condición (F).

"puede satisfacerse de hecho meramente porque el lenguaje con el que nos las vemos no posee un acopio suficiente de constantes extra-lógicas. La condición (F) podría ser considerada como suficiente para que la oración *X* se siga del conjunto *K* sólo si los nombres de todos los objetos posibles aparecieran en el lenguaje en cuestión. Esta suposición, sin embargo, es quimérica y nunca puede darse" (Tarski, 1936: 415-416).

Es decir, Tarski observa que para que un argumento sea un caso de consecuencia lógica no tiene por qué ser suficiente que todos los argumentos de la misma forma sean argumentos donde no es el caso que las premisas sean verdaderas y la conclusión falsa. Es concebible que sea posible interpretar las constantes no lógicas del argumento por medio de ciertos objetos (individuos, propiedades, etc.), de manera que las premisas así reinterpretadas sean verdaderas y la conclusión así reinterpretada sea falsa, y que sin embargo (algunos de) esos objetos no estén denotados por constantes no lógicas del lenguaje que está siendo considerado; en tal caso no diríamos que el argumento es un caso de consecuencia lógica, a pesar de que satisfaría la condición (F).

Para poner un ejemplo, supongamos que el lenguaje que estamos considerando es LAr^{+}. Puesto que tanto la relación "ser menor que" como la relación "ser predecesor inmediato de" son irreflexivas en el dominio de los números naturales, la oración '$\forall x(Nx \rightarrow \neg Mxx)$' sería declarada consecuencia lógica de todo conjunto de premisas por el criterio (F): ninguna sustitución de las constantes no lógicas 'N' y 'M' por otras constantes no lógicas de LAr^{+} convierte esa oración en otra falsa. Pero claramente '$\forall x(Nx \rightarrow \neg Mxx)$' no es una consecuencia lógica de, digamos, 'N0'. Esto se puede justificar, verbigracia, manteniendo fija la interpretación usual de '0' y 'N', pero observando que 'M' puede interpretarse por medio de la relación reflexiva "ser menor o igual que"; bajo esta interpretación, '$\forall x(Nx \rightarrow \neg Mxx)$' es falsa, a pesar de que 'N0' es verdadera. (La observación de Tarski según la cual es quimérico suponer que todos los objetos tengan nombre en el lenguaje puede justificarse, por ejemplo, observando que hay una cantidad no numerable de conjuntos de números naturales, pero, en los lenguajes considerados por él, sólo hay una cantidad numerable de constantes.)

La propuesta teórica de Tarski consistirá en ampliar el requisito expresado por la condición (F) de forma que se incorpore la idea de que un argumento

lógicamente correcto no puede ser *reinterpretado* haciéndose verdaderas las premisas y falsa la conclusión (y no meramente la idea de que no puede ser *convertido* en un argumento con premisas verdaderas y conclusión falsa sustituyendo constantes por constantes). En otras palabras, la idea de que una oración *X* es consecuencia lógica de un conjunto de oraciones *K* cuando *toda interpretación en que todas las oraciones de K son verdaderas es una interpretación en que X es verdadera*. (O, como a veces se dice, cuando toda interpretación *preserva* la verdad de las premisas en la conclusión.) Es a través de la delimitación de esta idea que la teoría de Tarski es susceptible de incorporar el componente modal de la relación intuitiva de consecuencia lógica.

Como dice Tarski, la idea de entender la noción de consecuencia lógica por medio de la noción de preservación de la verdad por toda interpretación no es una idea original suya, sino implícita en la práctica lógica y matemática de su tiempo (especialmente la de los matemáticos y lógicos interesados en ofrecer pruebas de independencia). Pero lo novedoso es que Tarski propone hacer precisa esta idea empleando el aparato desarrollado por él para la definición de conceptos semánticos, como los de satisfacción, definibilidad y verdad. Veamos cómo procede Tarski.

Tarski usa una cierta noción precisa de interpretación de un lenguaje formalizado. En nuestro ejemplo, una *interpretación* de *LAr* es una secuencia que asigna objetos apropiados a las constantes no lógicas de *LAr*: un objeto singular a '0', un conjunto de objetos singulares a 'N' y una relación binaria entre objetos singulares a 'M'. Además de la noción de interpretación, Tarski introduce la noción de *función oracional*. Una función oracional *O'* de una oración *O* es el resultado de sustituir las constantes no lógicas que aparecen en *O* de una manera uniforme por variables correspondientes de los tipos apropiados (y diferentes de las variables ya existentes en el lenguaje). Por ejemplo, la función oracional determinada por la oración '$\forall x(Nx \rightarrow \neg Mx0)$' sería la expresión '$\forall x(Px \rightarrow \neg Yxy)$' (en la cual 'P', 'Y' e 'y' son variables nuevas). Igualmente es posible definir la noción más general de *función formular*, de forma análoga, salvo que ahora *O* puede ser una fórmula abierta. Las funciones oracionales de oraciones de *LAr* no serán en general oraciones y, por tanto, no verificarán siempre la propiedad de ser verdaderas o falsas. Pero, en cambio, sí serán siempre verdaderas o falsas *con respecto a interpretaciones de LAr*; o, como Tarski dice, serán *satisfechas* o no por interpretaciones de *LAr*.

El concepto de satisfacción de una función oracional por una interpretación puede definirse usando los métodos tarskianos para la definición de conceptos semánticos. Tarski no da un ejemplo pero nosotros veremos el caso de *LAr*. Decimos que *la interpretación* $<a, A, R>$ *satisface la función formular X con respecto a una secuencia f* (que asigne valores a las variables originales de *LAr*) si y sólo si:

(α) (i) X es $\ulcorner Px_n \urcorner$[1] (para algún n) y $f(x_n) \in A$; o X es 'Py' y $a \in A$; o
(ii) X es $\ulcorner Yx_n x_m \urcorner$ (para algunos m y n) y $<f(x_n), f(x_m)> \in R$; o X es $\ulcorner Yyx_n \urcorner$ (para algún n) y $<a, f(x_n)> \in R$; o X es $\ulcorner Yx_n y \urcorner$ (para algún n) y $<f(x_n), a> \in R$; o X es 'Yyy' y $<a,a> \in R$; o
(iii) X es $\ulcorner x_n = x_m \urcorner$ (para algunos m y n) y $f(x_n) = f(x_m)$; o X es $\ulcorner y = x_n \urcorner$ (para algún n) y $a = f(x_n)$; o X es $\ulcorner x_n = y \urcorner$ (para algún n) y $f(x_n) = a$; o X es 'y=y'; o

(β) hay una función formular Y tal que X es $\ulcorner \neg Y \urcorner$ y $<a, A, R>$ no satisface Y con respecto a la secuencia f; o

(γ) hay funciones formulares Y y Z tales que X es $\ulcorner (Y \rightarrow Z) \urcorner$ y o bien $<a, A, R>$ no satisface Y con respecto a la secuencia f o $<a, A, R>$ satisface Z con respecto a la secuencia f; o, por último,

(δ) hay una función formular Y y un número n tales que X es $\ulcorner \forall x_n Y \urcorner$ y toda secuencia g que asigna valores a las variables (originales) de *LAr* y que difiere de f a lo sumo en lo que asigna a x_n es tal que $<a, A, R>$ satisface Y con respecto a g.

Es esta una definición recursiva, enteramente paralela a la definición de satisfacción de fórmulas por secuencias en el artículo de Tarski "El concepto de verdad en los lenguajes formalizados" (véase Tarski, 1935: 193; hemos procurado usar la misma terminología que en este artículo clásico de Tarski). De la misma manera que esa definición, la que acabamos de dar puede hacerse también explícita de forma enteramente análoga.

En términos de la noción definida, la noción de satisfacción de una función oracional es fácil de caracterizar. Decimos que *la interpretación* $<a, A, R>$

1. Sobre el uso de los símbolos '⌜' y '⌝' (las semicomillas) véase el Apéndice.

satisface la función oracional X si y sólo si $<a, A, R>$ satisface la función formular X con respecto a toda secuencia f que asigna valores a las variables de *LAr*. Esta definición es análoga a la definición tarskiana de verdad (véase Tarski, 1935: 195).

En términos de la noción definida de satisfacción, Tarski introduce la noción de *modelo* de una oración. Un modelo de una oración O es una interpretación del lenguaje de O que satisface la función oracional O' determinada por O; de forma más general, un modelo de un conjunto de oraciones K es una interpretación del lenguaje de las oraciones de K que satisface todas las funciones oracionales determinadas por oraciones de K. Y, en términos de la noción definida de modelo, Tarski propone una noción definida de consecuencia lógica: la oración X es una (consecuencia lógica)$_t$ de las oraciones del conjunto K si y sólo si todo modelo del conjunto K es también un modelo de la oración X (véase Tarski, 1936: 417). Tarski también propone una noción definida de verdad lógica (él usa la expresión 'verdad analítica') usando el mismo aparato: una oración O es una (verdad lógica)$_t$ si y sólo si toda interpretación del lenguaje de O es un modelo de O. (El subíndice 't' lo uso para poner énfasis en el hecho de que '(consecuencia lógica)$_t$' y '(verdad lógica)$_t$' denotan una relación y una propiedad *definidas* (por Tarski, de ahí la 't'), no la relación y la propiedad intuitivas.) Nociones análogas de (consecuencia lógica)$_t$ y (verdad lógica)$_t$ se pueden definir para otros lenguajes usando el mismo método seguido por nosotros con *LAr*, simplemente haciendo los cambios obvios.

V. Teoría, definición y corrección

¿En qué sentido es la definición ofrecida por Tarski una *teoría* de la noción intuitiva e importante de consecuencia lógica? La noción definida por Tarski, la noción de preservación de verdad por toda interpretación, *no* es la noción intuitiva de consecuencia lógica: cuando nos preguntamos si una conclusión se sigue o no de un conjunto de premisas dado, nuestra mente no se hace una pregunta acerca de secuencias, funciones formulares, conjuntos... Todas estas son nociones que aparecen en la noción definida por Tarski, pero nadie se pregunta por verdades acerca de ellas cuando se pregunta por verdades acerca de la noción *intuitiva* de consecuencia lógica. Si el único objetivo de una teoría de la consecuencia lógica es explicar el significado intuitivo de esa noción, entonces parece claro, sin necesidad de reflexionar mucho, que la definición de Tarski, si es que es una teoría de la consecuencia lógica, no es una buena teoría.

Pero una teoría filosófica de una noción intuitiva como la de consecuencia lógica no tiene por qué tener como único objetivo la explicación de su significado intuitivo. El principal objetivo de Tarski al proponer su definición no es éste, sino el objetivo de *caracterizar extensionalmente* la noción intuitiva de consecuencia lógica, en términos de un aparato de conceptos mejor comprendidos que el intuitivo de consecuencia lógica. Consideremos la noción intuitiva de "arco iris". Esta noción tiene un significado intuitivo para la gran mayoría de la gente; prácticamente todo el mundo puede asociar ideas intuitivas con ella,

pero la cantidad de gente familiarizada con la base óptica y física en general del arco iris es relativamente escasa. Supongamos que el físico *define* "arco iris" por medio de una serie de conceptos físicos técnicos, por ejemplo "arco de colores prismáticos causado por la refracción y reflexión de los rayos del sol en las gotas de lluvia". Con esta definición, el físico no intenta explicar el significado de la noción intuitiva de arco iris, sino, en parte, ofrecer una caracterización en términos menos intuitivos, mejor comprendidos, que se aplique *exactamente* a los arcos iris (y que de este modo explique o ilumine el hecho de que haya arcos iris). (Podría disputarse que un objetivo del físico o del científico en general sea asegurar la coextensionalidad de conceptos intuitivos y conceptos técnicos. En muchos casos, la investigación puede revelar que el concepto intuitivo es problemático, quizá incluso contradictorio, y, por tanto, su sustitución por un concepto técnico coextensional es ociosa. Pero en muchos otros casos la ciencia parece hacer justamente caracterizaciones extensionales de conceptos intuitivos en términos de conceptos técnicos; un ejemplo famoso en la bibliografía filosófica es la identidad extensional entre agua y H_2O.)

Como otras definiciones científicas, la definición de consecuencia lógica de Tarski tiene como uno de sus objetivos la caracterización extensional de una noción intuitiva en términos de conceptos técnicos, mejor comprendidos que el concepto intuitivo de consecuencia lógica. Entre estos conceptos técnicos vimos que se hallaban los de secuencia, función formular, conjunto y otros, de los que Tarski propone que tenemos una mejor comprensión –suministrada por la lógica y la matemática– que del concepto de consecuencia lógica. Esta mejor comprensión se muestra, en parte, en el hecho de que hay teorías muy útiles y bien estudiadas acerca de esos conceptos: teorías de la sintaxis de los lenguajes formales y teorías matemáticas de los conjuntos. Una de las esperanzas de quien ofrece una definición como la de Tarski es que nuestros conocimientos de esas teorías puedan aplicarse de modo fructífero en la obtención de resultados acerca de la relación *definida* de consecuencia lógica y, de este modo, si es que efectivamente son coextensionales, acerca de la noción *intuitiva* de consecuencia lógica.

¿Son coextensionales la noción de (consecuencia lógica)$_t$ y la noción intuitiva de consecuencia lógica? El problema de si son coextensionales es el problema de la corrección de la definición de (consecuencia lógica)$_t$: la definición será correcta en el caso de que todo argumento (de cualquier lenguaje formal) que sea

un ejemplo de consecuencia lógica en el sentido intuitivo lo sea también de (consecuencia lógica)$_t$, y viceversa. La primera de las dos partes de este problema la constituye la cuestión de si cuando una conclusión X es consecuencia lógica de un conjunto de premisas K ocurre también que todo modelo del conjunto K es un modelo de la oración X. Nuestras intuiciones acerca del concepto preteórico de consecuencia lógica, a pesar de ser precarias, nos permiten zanjar esta cuestión afirmativamente, a través del argumento siguiente.

Supongamos que X es consecuencia lógica de K pero (en busca de una contradicción) que hay una interpretación I que es modelo de K pero que no es modelo de X. Recordemos que es parte integral del concepto preteórico de consecuencia lógica que es un concepto formal y modal. Lo primero quiere decir que todo argumento *existente* de la misma forma que un argumento lógicamente correcto de K a X es también un ejemplo de consecuencia lógica; pero nuestras intuiciones también nos dicen que incluso si ampliamos nuestro lenguaje con nuevo vocabulario y de este modo introducimos *nuevos argumentos* de la misma forma que el de K a X, estos nuevos argumentos también han de ser lógicamente correctos si el original de hecho lo era. Así, pues, si introducimos nombres para las entidades asignadas por I a las constantes no lógicas de K y X, y se usan esos nombres para reemplazar a las constantes que les corresponden, obtenemos un argumento con conjunto de premisas K' y conclusión X' que es lógicamente correcto. Ahora bien, que el concepto de consecuencia lógica es modal quiere decir que cuando una conclusión X' es consecuencia lógica de un conjunto de premisas K', X' es implicada por necesidad lógica por las oraciones de K'. Pero si X' es implicada por necesidad lógica por las oraciones de K', entonces no es posible que todas las oraciones de K' sean verdaderas y X' falsa. Sin embargo, que I es modelo de K y no de X es lo mismo que decir que las oraciones de K' son verdaderas y X' es falsa. Esta es la contradicción que buscábamos para establecer la tesis de que todo argumento lógicamente correcto en el sentido intuitivo es también un caso de (consecuencia lógica)$_t$.

La otra parte del problema de la corrección la constituye la cuestión de si todo caso de (consecuencia lógica)$_t$ es un caso de consecuencia lógica en el sentido intuitivo. Una respuesta positiva habría de basarse *al menos* en dos argumentos: (1) un argumento para mostrar que la relación de (consecuencia lógica)$_t$ es una relación *formal*, es decir, para mostrar que (para cualesquiera K, K', X y X') si todo modelo de K es un modelo de X y el argumento de K a X tiene la misma

forma que el de K' a X', entonces todo modelo de K' es un modelo de X'; y (2) un argumento para mostrar que ésa relación es una relación *modal*, es decir, para mostrar que (para cualesquiera K y X) si todo modelo de K es un modelo de X, entonces K implica X por necesidad lógica. Decimos "al menos" porque observamos que formalidad y modalidad son propiedades *necesarias* de los argumentos que intuitivamente son lógicamente correctos, pero no dijimos que la conjunción de ambas propiedades fuera *suficiente* para que un argumento sea lógicamente correcto (esta tesis nos parece plausible pero quizá no sea universalmente aceptada; por otro lado, la tesis de que formalidad y modalidad son propiedades necesarias de los argumentos lógicamente correctos ha sido negada por quienes admiten que argumentos como

Doña Pura es soltera

Doña Pura es una mujer no casada

son lógicamente correctos a pesar de que otros argumentos con la misma forma no lo son). Naturalmente, si se pudiera mostrar que la relación de (consecuencia lógica)$_t$ no es formal o no es modal, entonces la respuesta a la segunda parte del problema de la corrección sería negativa: habría argumentos lógicamente correctos en el sentido tarskiano, pero no lógicamente correctos en el sentido preteórico.

Tarski mismo aludió a un argumento que puede modificarse para mostrar que la relación de (consecuencia lógica)$_t$ es formal (véase Tarski, 1936: 417). El argumento insinuado por Tarski propone que si una oración X es consecuencia lógica de un conjunto de oraciones K en el sentido definido por él, entonces el argumento de K a X cumple la condición (F). El argumento es simple: supongamos que todo modelo de K es modelo de X, y (en busca de un absurdo) que hay un argumento de, digamos, K' a X' con la misma forma pero tal que todas las oraciones de K' son verdaderas y X' es falsa; entonces tomemos la interpretación de las constantes no lógicas de K' y X' en la cual las oraciones de K' son verdaderas y X' es falsa, y usémosla para interpretar de la forma natural las constantes no lógicas de K y X; en esta interpretación, las oraciones de K son verdaderas y X es falsa, contra la suposición de que todo modelo de K es un modelo de X. (Este argumento ha sido ampliamente

malentendido en la bibliografía sobre el tema. Para una discusión de los malentendidos y de cómo corregirlos véase Gómez Torrente, 1998.)

Ahora bien, un examen atento revela que el argumento tarskiano sólo funciona bajo la suposición de que la interpretación, *en el sentido intuitivo* de 'interpretación', de las constantes no lógicas de K' y X' que hace verdaderas a las oraciones de K' y falsa a X' es (o puede transformarse de alguna manera apropiada en) una interpretación *en el sentido técnico* usado en la definición de Tarski y que igualmente haga verdaderas a las oraciones de K' y falsa a X' (véase Gómez Torrente, 1996a, n. 21; posiblemente la noción de interpretación usada por Tarski en 1936 está más cerca de la noción intuitiva que la noción teórica común hoy en día). Aunque esta suposición no está demostrada para el sentido técnico de 'interpretación' común hoy en día, es razonable hacerla (como se explica en el capítulo VIII), y vale sin duda para el caso de que K y X (y por tanto K' y X') pertenezcan a un lenguaje de primer orden (lo veremos en el capítulo VII). Por tanto, el argumento de Tarski muestra que la noción de (consecuencia lógica)$_t$ cumple la condición (F) para lenguajes de primer orden, y hace muy probable esta tesis para lenguajes de orden superior.

El argumento de Tarski puede modificarse fácilmente para mostrar que la relación de (consecuencia lógica)$_t$ es formal, en el sentido que nosotros le hemos dado a este término. Supongamos que todo modelo de K es modelo de X y (en busca de un absurdo) que hay un argumento de, digamos, K' a X' con la misma forma pero tal que hay un modelo de K' que no es un modelo de X'; tomemos la interpretación de las constantes no lógicas de K' y X', en la cual las oraciones de K' son verdaderas y X' es falsa, y usémosla para interpretar de la forma natural las constantes no lógicas de K y X; en esta interpretación, las oraciones de K son verdaderas y X es falsa, contra la suposición de que todo modelo de K es un modelo de X.

Aunque Tarski alude a la esencia de un razonamiento que muestra que la noción de (consecuencia lógica)$_t$ es formal, en vano buscaríamos en él una justificación de que esta noción definida es modal en nuestro sentido. Tarski era profundamente escéptico acerca de la inteligibilidad y el valor filosófico de las propiedades modales, como la necesidad (lógica o de cualquier otra especie), la aprioridad, la analiticidad, etc., y es dudoso que sintiera jamás inclinación a usar estas nociones en razonamientos sobre los que pretendiese descargar peso alguno. El principal argumento que Tarski parece haber tenido en mente, para

justificar la corrección de su definición de (consecuencia lógica)$_t$, es el argumento que hemos comentado hace tres párrafos.

Que Tarski no se preocupase por la cuestión no quiere decir que no sea una cuestión legítima para quien no comparta el escepticismo radical de Tarski sobre las nociones modales. El punto de vista que adoptamos aquí es que tales nociones, aunque oscuras, permiten en ocasiones el razonamiento intuitivo con ellas y el establecimiento de conclusiones interesantes expresadas en términos de ellas. Que lo permitan en el caso que nos concierne no es obviamente imposible. Se trata de una cuestión que merece la pena explorar.

Para ilustrar la forma en que el razonamiento con nociones modales podría utilizarse para sacar conclusiones de interés para la cuestión de la corrección de la definición de (consecuencia lógica)$_t$, discutiremos una objeción que puede hacerse a la definición. La objeción se basa en la siguiente observación. Oraciones de *LAr* sin cuantificadores relativizados, como '$\exists x\exists x.\neg x=x.$' o '$\exists x\exists x.\exists x..((\neg x=x.\wedge\neg x=x..)\wedge\neg x.=x..)$', cuyos significados intuitivos son "hay dos objetos diferentes" y "hay tres objetos diferentes", respectivamente, son (consecuencias lógicas)$_t$ de cualquier conjunto de premisas. En otras palabras, son (verdades lógicas)$_t$, ya que son verdaderas en toda interpretación de las constantes no lógicas de *LAr*. Esto es así porque esas oraciones *no contienen constantes no lógicas*; así, pues, vacuamente, si son verdaderas son verdaderas en toda interpretación; y son verdaderas (pues, naturalmente hay al menos tres objetos en el mundo), así que son verdaderas en toda interpretación. La objeción a la corrección de la definición de (consecuencia lógica)$_t$ es, entonces, que '$\exists x\exists x.\neg x=x.$' y '$\exists x\exists x.\exists x..((\neg x=x.\wedge\neg x=x..)\wedge\neg x.=x..)$' no son verdades lógicamente necesarias, implicadas por necesidad lógica por todo conjunto de premisas –pues no parece una imposibilidad lógica que el mundo hubiese tenido menos de dos objetos, ni siquiera que hubiese tenido menos de uno–, y, sin embargo, ambas oraciones son (consecuencias lógicas)$_t$ de todo conjunto de premisas. Si la objeción es justa, la definición de (consecuencia lógica)$_t$ será incorrecta: ciertos casos de (consecuencia lógica)$_t$ no serán casos de consecuencia lógica en el sentido intuitivo.

La primera observación pertinente sobre esta objeción es que Tarski, al menos en la época en que propuso su definición, aceptaba que las oraciones verdaderas sin constantes no lógicas que tienen implicaciones sobre la cardinalidad del conjunto de objetos del mundo son verdades lógicas intuitivas

y no habría visto un problema en el hecho de que sean verdades lógicas de acuerdo con su definición. (Pero, naturalmente, no habría admitido que oraciones como '$\exists x\exists x.\neg x=x.$' son verdades lógicas cuando el contexto dicta que no son más que *abreviaciones* de oraciones con cuantificadores relativizados a un predicado no lógico, por ejemplo, '$\exists x\exists x.((Nx\wedge Nx.)\wedge\neg x=x.)$'.) En esta convicción no estaba solo en aquel tiempo. Ramsey –siguiendo a Wittgenstein– propuso que las oraciones sin constantes no lógicas que tienen implicaciones sobre la cardinalidad del conjunto de objetos del mundo son tautologías si son verdaderas, y contradicciones si son falsas. Otros filósofos de tendencia logicista, como Carnap y Hempel, sostuvieron ideas parecidas. (Para una breve discusión y referencias véase Gómez Torrente, 1996.)

De todos modos, Tarski mismo observó que quizá el concepto definido más interesante de verdad lógica no era el de verdad en toda interpretación de las constantes no lógicas, sino el de verdad en toda interpretación de las constantes no lógicas *sacada de cualquier universo* (no vacío) *de objetos* (véase Tarski, 1935: 239-240). Oraciones como '$\exists x\exists x.\neg x=x.$' no caen bajo la extensión de este concepto, que es también definible usando el aparato tarskiano (de hecho, la definición consiste en una modificación minúscula de la definición de (consecuencia lógica)$_t$). En el siguiente capítulo veremos de una forma más precisa cómo se define, y por qué '$\exists x\exists x.\neg x=x.$' no cae bajo su extensión. Además, también veremos que ambas nociones definidas pueden verse como equivalentes dada una cierta suposición, inusual pero razonable, acerca de la forma implícita de las oraciones de los lenguajes formales. La conclusión será que la objeción que hemos expuesto no es una objeción válida a ninguna de las dos definiciones.

VI. La definición tarskiana usual de consecuencia lógica

Usemos de nuevo el lenguaje *LAr* como ejemplo. Digamos que una *estructura* para *LAr* es una secuencia que tiene como primer miembro a un conjunto no vacío U y que asigna objetos "sacados" de U a las constantes no lógicas de *LAr*: un objeto singular de U a '0', un conjunto de objetos de U a 'N' y una relación binaria entre objetos de U a 'M'.

Definimos entonces el concepto de satisfacción de una función oracional por una estructura. Decimos que *la estructura* $<U, a, A, R>$ *satisface la función formular* X *con respecto a una secuencia* f (que asigne valores en U a las variables de *LAr*) si y sólo si:

(α) (i) X es $\ulcorner Px_n \urcorner$ (para algún n) y $f(x_n) \in A$; o X es 'Py' y $a \in A$; o
(ii) X es $\ulcorner Yx_n x_m \urcorner$ (para algunos m y n) y $<f(x_n), f(x_m)> \in R$; o X es $\ulcorner Yyx_n \urcorner$ (para algún n) y $<a, f(x_n)> \in R$; o X es $\ulcorner Yx_n y \urcorner$ (para algún n) y $<f(x_n), a> \in R$; o X es 'Yyy' y $<a, a> \in R$; o
(iii) X es $\ulcorner x_n = x_m \urcorner$ (para algunos m y n) y $f(x_n) = f(x_m)$; o X es $\ulcorner y = x_n \urcorner$ (para algún n) y $a = f(x_n)$; o X es $\ulcorner x_n = y \urcorner$ (para algún n) y $f(x_n) = a$; o X es 'y=y'; o

(β) hay una función formular Y tal que X es $\ulcorner \neg Y \urcorner$ y $<U, a, A, R>$ no satisface Y con respecto a la secuencia f; o

(γ) hay funciones formulares Y y Z tales que X es $\ulcorner (Y \rightarrow Z) \urcorner$ y o bien $<U, a, A, R>$ no satisface Y con respecto a la secuencia f o $<U, a, A, R>$ satisface Z con respecto a la secuencia f; o, por último,

(δ) hay una función formular Y y un número n tales que X es $\ulcorner\forall x_n Y\urcorner$ y toda secuencia g que asigna valores en U a las variables de LAr y que difiere de f a lo sumo en lo que asigna a x_n es tal que $<U, a, A, R>$ satisface Y con respecto a g.

Esta definición es enteramente análoga a la definición de satisfacción de una función formular por una interpretación con respecto a una secuencia. Únicamente se hace la modificación de restringir el recorrido de las secuencias (de valores de las variables de LAr) al universo U de la estructura dada.

Podemos decir, entonces, que *la estructura* $<U, a, A, R>$ *satisface la función oracional* X si y sólo si $<U, a, A, R>$ satisface la función formular X con respecto a toda secuencia f que asigna valores en U a las variables de LAr. Siguiendo con las definiciones análogas, digamos que una *estructura modelo* de una oración O es una estructura para el lenguaje de O que satisface la función oracional O' determinada por O; de forma más general, una estructura modelo de un conjunto de oraciones K es una estructura para el lenguaje de las oraciones de K que satisface todas las funciones oracionales determinadas por oraciones de K. Finalmente, en términos de este aparato de nociones podemos ofrecer la definición de nuevas nociones de consecuencia y verdad lógicas:

(CLT) Una oración O es una (consecuencia lógica)$_T$ de un conjunto de oraciones K si y sólo si toda estructura modelo del conjunto K es también una estructura modelo de la oración O.

(VLT) Una oración O es una (verdad lógica)$_T$ si y sólo si toda estructura para el lenguaje de O es una estructura modelo de O.

'CLT' y 'VLT' significan, respectivamente, "consecuencia lógica tarskiana" y "verdad lógica tarskiana". (El subíndice 'T' lo uso de nuevo para poner énfasis en el hecho de que '(consecuencia lógica)$_T$' y '(verdad lógica)$_T$' denotan una relación y una propiedad definidas, no la relación y la propiedad intuitivas.) De nuevo, nociones análogas de (consecuencia lógica)$_T$ y (verdad lógica)$_T$ se pueden definir para otros lenguajes usando el mismo método.

Hay varias razones para que hablemos de estas nociones como nociones "tarskianas". Una es que, como se ve claramente en nuestra presentación, ellas

y las nociones definidas por Tarski que examinamos en el capítulo IV son definibles usando métodos enteramente tarskianos, se obtienen la una de la otra a través de modificaciones minúsculas, y ambas son mencionadas por Tarski. Otra razón es que las nociones de "(consecuencia lógica)$_T$" y "(verdad lógica)$_T$" son las nociones más comúnmente atribuidas a Tarski en la bibliografía lógica y filosófica (nosotros seguiremos la práctica de atribuírselas, pues, como hemos visto, está plenamente justificada). Una tercera, más importante, es que las diferencias entre ambos tipos de nociones son, como explicaremos en breve, más aparentes que reales.

'$\exists x\exists x.\neg x=x.$' no es (consecuencia lógica)$_T$ de todo conjunto de premisas, ni, por tanto, (verdad lógica)$_T$. La razón es que hay estructuras para LAr que no son modelos de '$\exists x\exists x.\neg x=x.$', a pesar de que vacuamente son modelos del conjunto vacío de premisas. Una de ellas es la estructura $<U, a, A, R>$ donde U es el conjunto $\{0\}$ (el conjunto cuyo único miembro es el número cero), a es 0, A es $\{0\}$ y R es la relación $\{<0,0>\}$. Recordemos que '$\exists x\exists x.\neg x=x.$' es una abreviación definicional de '$\neg\forall x\forall x.x=x.$' (o, para ser completamente estrictos, de una oración trivialmente equivalente a ésta: '$\neg\forall x\neg\neg\forall x.\neg\neg x=x.$'; pero ignoramos esta complicación). La función formular y oracional, determinada por '$\neg\forall x\forall x.x=x.$' es la oración misma. Por tanto, $<U, a, A, R>$ es modelo de '$\neg\forall x\forall x.x=x.$' si y sólo si $<U, a, A, R>$ satisface la función formular '$\neg\forall x\forall x.x=x.$' con respecto a toda secuencia f que asigna valores en U a las variables de LAr. Ahora bien, la única secuencia que asigna valores en U a las variables de LAr asigna 0 a todas ellas (pues no hay otro objeto en U); llamémosla h. Y $<U, a, A, R>$ no satisface '$\neg\forall x\forall x.x=x.$' con respecto a h. Veámoslo.

Hay una función formular X tal que '$\neg\forall x\forall x.x=x.$' es $\ulcorner\neg X\urcorner$, a saber '$\forall x\forall x.x=x.$'. Por tanto, vayamos al caso (β). Supongamos, en busca de un absurdo, que $<U, a, A, R>$ satisface '$\neg\forall x\forall x.x=x.$' con respecto a h. Si fuera así, $<U, a, A, R>$ no satisfaría '$\forall x\forall x.x=x.$' con respecto a h. Vayamos al caso (δ). Si $<U, a, A, R>$ no satisface '$\forall x\forall x.x=x.$' con respecto a h ha de haber una secuencia g que asigna valores en U a las variables de LAr y que difiere de h a lo sumo en lo que asigna a 'x' y tal que $<U, a, A, R>$ no satisface '$\forall x.x=x.$' con respecto a g. Pero entonces g ha de ser igual a h. Así, $<U, a, A, R>$ no satisface '$\forall x.x=x.$' con respecto a h. Volvamos a (δ). Como $<U, a, A, R>$ no satisface '$\forall x.x=x.$' con respecto a h, ha de haber

una secuencia f que asigna valores en U a las variables de LAr y que difiere de h a lo sumo en lo que asigna a 'x$_n$' y tal que $<U, a, A, R>$ no satisface 'x=x$_n$' con respecto a f. Pero de nuevo f ha de ser h, así que $<U, a, A, R>$ no satisface 'x=x$_n$' con respecto a h. En virtud del caso (α)(iii), esto implica que h('x') no es igual a h('x$_n$'). Pero éste es el absurdo que buscábamos, ya que, como observamos, h('x')=h('x$_n$')=0.

Así pues, '∃x∃x$_n$¬x=x$_n$' no es (consecuencia lógica)$_T$ de todo conjunto de premisas (pues no lo es del conjunto vacío), ni, consiguientemente, (verdad lógica)$_T$. De forma análoga se ve que oraciones como '∃x∃x∃x$_n$((¬x=x∀¬x=x$_n$) ∀¬x$_n$=x$_n$)' no son (verdades lógicas)$_T$. Por tanto, la objeción que vimos a la corrección de la definición de "(consecuencia lógica)$_t$" al final del capítulo anterior no es una objeción válida a la corrección de la definición de "(consecuencia lógica)$_T$". Pero, además, ahora que hemos visto cómo definir "(consecuencia lógica)$_T$", es posible ver también que la definición de "(consecuencia lógica)$_t$" sería también inmune a la objeción discutida si simplemente se hiciera una suposición inusual pero razonable acerca de la forma implícita de las oraciones de los lenguajes formalizados para los que está diseñada.

Esta suposición consistiría en entender que cuantificadores aparentemente no relativizados como '∃x' o '∃x$_n$' están de hecho relativizados a través de una construcción sintáctica especial omitida a un predicado *no lógico* 'U', cuya interpretación pone límites a las interpretaciones de los otros predicados no lógicos –que han de "sacarse" de la interpretación de 'U'–. En el caso de LAr, a las constantes no lógicas manifiestas del lenguaje habría que añadir la constante no lógica no manifiesta 'U', y una construcción sintáctica especial para la relativización de todo cuantificador a 'U'; por ejemplo, una oración como '∃x∃x$_n$¬x=x$_n$' sería en realidad una forma de abreviar la oración '∃xUx∃x$_n$Ux$_n$¬x=x$_n$'. La noción apropiada de *interpretación* de LAr coincidiría entonces extensionalmente con la noción de *estructura* para LAr: una *interpretación* de LAr es una secuencia que asigna objetos apropiados a las constantes no lógicas (manifiestas y no manifiestas) de LAr: un conjunto no vacío de objetos singulares a 'U', un objeto singular de U a '0', un conjunto de objetos de U a 'N' y una relación binaria entre objetos de U a 'M'.

Es posible usar esta noción de interpretación en definiciones de nociones apropiadas de (1) satisfacción de una función formular por una interpretación $<U, a, A, R>$ con respecto a una secuencia de objetos de U, (2) satisfacción

de una función oracional por una interpretación $<U, a, A, R>$, (3) interpretación modelo de una oración y de un conjunto de oraciones, y (4) (consecuencia lógica)$_t$ y (verdad lógica)$_t$. (Las nociones de (4) no serían diferentes a las nociones definidas por Tarski que vimos en el capítulo IV, sino más bien versiones diferentes de una misma noción, en una de las cuales se toma a 'U' como constante lógica y en la otra no.) El resultado es que, como se comprueba fácilmente, las nociones de (consecuencia lógica)$_t$ y (verdad lógica)$_t$ así definidas son extensionalmente equivalentes a las de (consecuencia lógica)$_T$ y (verdad lógica)$_T$ definidas en este capítulo. En particular, '∃x∃x$_n$¬x=x$_n$' no es ni (verdad lógica)$_t$ ni (verdad lógica)$_T$. Que no es (verdad lógica)$_T$ lo hemos visto ya: hay un universo U y una asignación de objetos sacados de U a las constantes no lógicas manifiestas de LAr que forman una estructura $<U, a, A, R>$ que no satisface '∃x∃x$_n$¬x=x$_n$'. Que no es (verdad lógica)$_t$ se puede justificar usando esa misma estructura, vista ahora como *interpretación* de las constantes no lógicas (manifiestas y no manifiestas) de LAr; '∃x∃x$_n$¬x=x$_n$' es sólo una abreviación de '∃xUx∃x$_n$Ux$_n$¬x=x$_n$', y $<U, a, A, R>$ no es un modelo de esta oración, pues no hay dos objetos de U que sean diferentes.

La objeción que enunciamos al final del capítulo V no es una objeción válida a la corrección de las definiciones de (consecuencia lógica)$_t$ y (consecuencia lógica)$_T$, supuesto que estemos dispuestos a entender la primera de ellas bajo la suposición razonable hecha explícita en este capítulo. Pero, además, mostrando que no es válida, hemos mostrado de paso que un escepticismo radical sobre la posibilidad del razonamiento intuitivo con nociones modales como la de necesidad lógica es una postura injustificada. En particular, es posible establecer conclusiones acerca de relaciones existentes entre la noción informal de necesidad lógica y las nociones técnicas definidas de (consecuencia lógica)$_t$ y (consecuencia lógica)$_T$. A saber, nuestra conclusión es que ciertas oraciones "lógicamente contingentes", verdaderas pero que no todo conjunto de premisas implica por necesidad lógica, no son (consecuencia lógica)$_t$ ni (consecuencia lógica)$_T$ de todo conjunto de premisas.

Esta conclusión es favorable a las definiciones tarskianas, pero es de un alcance muy limitado. Lo único que establece es que un supuesto contraejemplo a la tesis de que todo caso de (consecuencia lógica)$_t$ (o de (consecuencia lógica)$_T$) es un caso de implicación por necesidad lógica no es tal contraejemplo. ¿Es posible dar una justificación de la *tesis general*? Es decir, ¿es posible ofrecer un

razonamiento para mostrar que (para cualquier lenguaje L, cualquier conjunto de oraciones K de L y cualquier oración X de L) si todo modelo (o estructura modelo) de K es un modelo (o estructura modelo) de X entonces K implica X por necesidad lógica? Es ésta una cuestión harto complicada. La principal complicación proviene del hecho de que aunque tenemos un conocimiento de la noción de implicación por necesidad lógica que nos permite evaluar con certidumbre muchos casos particulares, nuestro conocimiento es algo difuso y, quizá, insuficiente para zanjar una cuestión tan general. Una analogía tal vez sea de utilidad aquí. Lo que yo sé sobre la noción "ser un querubín" es algo difuso. Sé lo suficiente como para decir con convicción que San Rafael no es un querubín, al ser un arcángel y ser el grado de los querubines disjunto del de los arcángeles; pero no sé lo suficiente como para afirmar con convicción la tesis general de que ningún ángel con nombre conocido es un querubín (en el momento en que escribo no sé, por ejemplo, si Lucifer es un querubín o no).

En los próximos tres capítulos intentaremos ahondar en nuestras intuiciones sobre la noción informal de implicación por necesidad lógica. Haremos diversas hipótesis sobre su contenido en términos de las cuales investigaremos (la segunda parte de) el problema de la corrección de las definiciones tarskianas de consecuencia lógica. Pero diremos ya que, a partir de ahora, nos concentraremos exclusivamente en la noción tarskiana usual, es decir, la noción de (consecuencia lógica)$_T$, dada por medio de la definición (CLT). Como hemos visto, esto no tiene mayor trascendencia, pues, bajo una suposición razonable, esta noción es equivalente a la noción originalmente definida por Tarski. Además, al versar sobre la noción usual, nuestra exposición y el significado de las conclusiones a que lleguemos resultarán así más claros al lector que ya esté familiarizado con ella a través de libros o cursos de lógica.

VII. CONSECUENCIA LÓGICA Y GENERALIDAD

Un importante grupo de filósofos que han reflexionado sobre el componente modal en el concepto de consecuencia lógica han procurado iluminarlo o explicarlo a través de las relaciones entre la noción de necesidad y la noción de *generalidad*. La tesis más extrema sobre estas relaciones es la de algunos filósofos según los cuales una grandísima parte de (o quizá todas) las atribuciones de la propiedad de necesidad que hacemos normalmente son, en realidad, ocultas atribuciones de generalidad acerca del mundo real. Por ejemplo, cuando decimos que necesariamente si Juan se come el helado de la nevera entonces se disgustará doña Pura, lo que hacemos es simplemente decir que en *todas* las ocasiones (actuales) en que alguien se come algo de la nevera de doña Pura, ésta se disgusta. Otro ejemplo: cuando decimos que las leyes de la física son necesarias, lo que hacemos es simplemente decir que carecen de excepciones (en el mundo actual), que *siempre* se aplican con verdad.

En el caso de la noción de implicación por necesidad lógica, los filósofos que la han intentado explicar por medio de la noción de generalidad han supuesto que tal implicación se da entre un conjunto de premisas K y una conclusión X cuando hay un conjunto de expresiones de K y X tal que *para todo modo de interpretar* esas expresiones de forma que todas las oraciones de K sean verdaderas, también X será verdadera. Naturalmente, las formas de entender "modo de interpretar" serán posiblemente diferentes en los distintos filósofos que sostengan este punto de vista. Bolzano lo sostuvo y, en su caso,

un "modo de interpretar" una proposición es simplemente una asignación de *ideas* o *conceptos* a los conceptos que se toman como reinterpretables y que entran en la composición de la proposición (nótese que Bolzano no habla de oraciones y expresiones, sino de proposiciones y conceptos). Bolzano también propuso que la noción de consecuencia lógica es simplemente una parte de la noción general de implicación por necesidad lógica tal como él la entendía: la noción de consecuencia lógica es la noción de preservación general de la verdad por todo modo de reinterpretar un cierto conjunto de conceptos, los conceptos que Bolzano llamó *no lógicos* (véase el § 155 de su *Wissenschaftslehre*). (Bolzano también sostuvo que Aristóteles había entendido la necesidad propia de las implicaciones silogísticas correctas de una forma similar a la suya, si bien esta opinión de Bolzano no sería universalmente aceptada.)

La teoría de la consecuencia lógica de Tarski se basa en un punto de vista estrechamente emparentado con el de Bolzano (las conexiones históricas entre ambos autores –que pueden haberse generado a través de relaciones maestro-discípulo– están aún por explorar). Las definiciones tarskianas proponen la coextensionalidad del concepto de consecuencia lógica y el concepto de preservación de la verdad por *toda* estructura, en cierto sentido preciso de "estructura". Sin embargo, como dijimos ya, en Tarski esta teoría no va acompañada de ninguna tesis sustancial respecto a las relaciones entre la noción intuitiva de implicación por necesidad lógica y la noción de generalidad.

Si es posible explicar la noción de implicación por necesidad lógica en términos de una cuantificación universal sobre "modos de interpretar" oraciones, y aceptamos en la misma línea que la noción de consecuencia lógica es explicable en términos de una cuantificación universal sobre "modos de interpretar" las constantes no lógicas, ¿qué conclusiones es posible sacar que sean pertinentes para el problema de la corrección de (CLT)? La respuesta depende de cuál sea exactamente la noción de "modo de interpretar" que tengamos en mente. Nótese que si la noción que tenemos en mente de "modo de interpretar" las constantes no lógicas de un lenguaje es simplemente la noción precisa de *estructura* para ese lenguaje en que se apoya (CLT), entonces trivialmente siempre que X fuera una (consecuencia lógica)$_T$ de K, K tendría como consecuencia lógica a X, y, por tanto, K implicaría por necesidad lógica a X. Pero hay varias formas de entender "modo de interpretar" bajo las cuales no todo modo de interpretar las constantes no lógicas de un lenguaje

es una estructura para ese lenguaje. Para estas formas de entender la noción, la tesis mencionada no es trivial. En este capítulo hablaremos muy someramente de tres de estas formas de entender "modo de interpretar", muy relacionadas entre sí y con la noción de estructura; prestaremos especial atención a una de ellas.

Dijimos que el primer componente de una estructura era siempre un *conjunto* no vacío. En ese momento no fuimos muy precisos sobre la noción de conjunto, porque no era necesario serlo. Pero ahora nos conviene llamar la atención sobre el hecho de que, en los libros de lógica donde se presenta la definición (CLT), la noción de conjunto por medio de la cual se define la noción de estructura es una noción matemática técnica. Los conjuntos son un cierto tipo de colecciones de objetos, de las que se ocupa la teoría matemática de conjuntos. Muchas colecciones especificables o definibles de objetos son conjuntos en el sentido de la teoría matemática de los conjuntos: en número ascendente de elementos, la colección de los poemas de Quevedo, la colección de los granos de arena de las playas de Argentina, la colección de los números naturales, la colección de los conjuntos de conjuntos de números naturales, y colecciones mucho mayores aún que éstas, son todas ejemplos de conjuntos (y conjuntos no vacíos); por ello, hay estructuras que los tienen como universos. Pero la teoría matemática de conjuntos dice también que ciertas colecciones no son conjuntos, son más grandes que cualquier conjunto. Por ejemplo, la colección de todos los conjuntos no es un conjunto. (Una prueba de esto es como sigue. Supongamos que C es la colección de todos los conjuntos y todos los objetos que no son colecciones (los "individuos"), y que C es ella misma un conjunto. Un axioma de la teoría de conjuntos dice que toda subcolección de un conjunto es un conjunto (una formulación en segundo orden del axioma de separación). Por este axioma, la colección $C' = \{c \in C : c \notin c\}$, que es una subcolección de C, es un conjunto. Pero entonces tanto la suposición de que $C' \in C'$ como la de que $C' \notin C'$ son contradictorias. Si $C' \in C'$, entonces claramente $C' \notin C'$; y si $C' \notin C'$, como además $C' \in C$, tenemos que $C' \in C'$. Así, hemos de negar nuestro supuesto, y C no es un conjunto.) Hay muchos otros ejemplos de colecciones que no son conjuntos.

A pesar de que la colección de todos los conjuntos y otras colecciones no son conjuntos, muchos lógicos han señalado que nada parece oponerse a la idea de considerar "modos de interpretar" lenguajes formalizados como *LAr*,

cuyos universos sean esas colecciones. Por ejemplo, parecería deseable ser capaz de considerar el siguiente "modo de interpretar" las constantes no lógicas del lenguaje *LAr*: como universo *U* tomamos la colección de todos los conjuntos y todos los individuos, a la constante individual '0' le asignamos el conjunto vacío (un elemento de *U*), al predicado 'N' le asignamos la colección de todos los conjuntos (una subcolección de *U*), y al predicado 'M' le asignamos la relación "ser elemento de" entre miembros de *U* (una colección de pares ordenados con miembros en *U*). Es deseable considerar modos como éste de interpretar las constantes no lógicas de *LAr*, simplemente porque así se han interpretado intuitivamente algunos lenguajes usados por los matemáticos para la formalización de la teoría de los conjuntos.

Digamos que una *clase* es una colección cuyos elementos son conjuntos o individuos. Poniendo de nuevo a *LAr* como ejemplo, digamos que un *(modo de interpretar)*$_{Cl}$ el lenguaje *LAr* es una secuencia $<U, a, A, R>$ donde *U* es una *clase* no vacía, *a* es un objeto elemento de *U*, *A* es una subclase de *U* y *R* es una clase de pares ordenados de elementos de *U*. Supongamos, además, que un conjunto de premisas *K* tiene como *(consecuencia lógica)*$_{Cl}$ a una conclusión *X* cuando para todo (modo de interpretar)$_{Cl}$ las oraciones de *K* de forma que todas ellas sean verdaderas, también *X* será verdadera. La pregunta que se plantea ante tal forma de entender la noción de consecuencia lógica es entonces si siempre que *X* es una (consecuencia lógica)$_{T}$ de *K*, *K* tiene como (consecuencia lógica)$_{Cl}$ a *X*. Dado que hay más (modos de interpretar)$_{Cl}$ que estructuras, la respuesta no es inmediata. Veremos en seguida una contestación parcial a esta pregunta; pero antes definamos otras dos formas de entender la noción de "modo de interpretar", y enunciemos las correspondientes nociones de consecuencia lógica a que dan lugar en un análisis generalista.

Una estructura es un modo de interpretar un lenguaje donde el universo es un conjunto de objetos del mundo real. Parece deseable considerar también modos de interpretar lenguajes cuyos universos sean conjuntos de objetos posibles, pues claramente muchos conjuntos de objetos posibles no son conjuntos de objetos reales. Ello da pie a la siguiente noción. Un *(modo de interpretar)*$_{CoP}$ el lenguaje *LAr* es una secuencia $<U, a, A, R>$, donde *U* es un conjunto no vacío de objetos posibles (todos los cuales existen juntos en un mundo posible *m*), *a* es un objeto elemento de *U*, *A* es un subconjunto de *U* y *R* es un conjunto de pares ordenados de elementos de *U*. Digamos que un conjunto de

premisas *K* tiene como *(consecuencia lógica)*$_{CoP}$ a una conclusión *X* cuando para todo (modo de interpretar)$_{CoP}$ las oraciones de *K* de forma que todas ellas sean verdaderas, también *X* será verdadera. La pregunta aquí es si siempre que *X* es una (consecuencia lógica)$_{T}$ de *K*, *K* tiene como (consecuencia lógica)$_{CoP}$ a *X*. De nuevo la respuesta no es inmediata, dado que hay más (modos de interpretar)$_{CoP}$ que estructuras.

Conjugando las ideas intuitivas que dan lugar a la formulación de las nociones de (modo de interpretar)$_{Cl}$ y de (modo de interpretar)$_{CoP}$, aparece otra noción de "modo de interpretar", aún más general: un modo de interpretar cuyo universo es una *clase posible*. Digamos que un *(modo de interpretar)*$_{ClP}$ el lenguaje *LAr* es una secuencia $<U, a, A, R>$, donde *U* es una clase no vacía de conjuntos e individuos posibles (todos los cuales existen juntos en un mundo posible *m*), *a* es un objeto elemento de *U*, *A* es una subclase de *U* y *R* es una clase de pares ordenados de elementos de *U*. Digamos que un conjunto de premisas *K* tiene como *(consecuencia lógica)*$_{ClP}$ a una conclusión *X* cuando para todo (modo de interpretar)$_{ClP}$ las oraciones de *K* de forma que todas ellas sean verdaderas, también *X* será verdadera. La pregunta correspondiente es si siempre que *X* es una (consecuencia lógica)$_{T}$ de *K*, *K* tiene como (consecuencia lógica)$_{ClP}$ a *X*. Y de nuevo la respuesta no es inmediata, pues naturalmente hay incluso más (modos de interpretar)$_{ClP}$ que (modos de interpretar)$_{CoP}$ y que (modos de interpretar)$_{Cl}$.

Resumiendo lo que sabemos de forma inmediata acerca de las nociones que hemos descrito: toda estructura es un (modo de interpretar)$_{Cl}$, un (modo de interpretar)$_{CoP}$ y un (modo de interpretar)$_{ClP}$; todo (modo de interpretar)$_{Cl}$ es un (modo de interpretar)$_{ClP}$; y todo (modo de interpretar)$_{CoP}$ es un (modo de interpretar)$_{ClP}$. Por tanto: si *K* tiene como (consecuencia lógica)$_{Cl}$ a *X*, entonces *K* tiene como (consecuencia lógica)$_{T}$ a *X*; si *K* tiene como (consecuencia lógica)$_{CoP}$ a *X*, entonces *K* tiene como (consecuencia lógica)$_{T}$ a *X*; si *K* tiene como (consecuencia lógica)$_{ClP}$ a *X*, *K* tiene como (consecuencia lógica)$_{T}$ a *X*; además: si *K* tiene como (consecuencia lógica)$_{ClP}$ a *X*, *K* tiene como (consecuencia lógica)$_{Cl}$ a *X*; y, finalmente, si *K* tiene como (consecuencia lógica)$_{ClP}$ a *X*, *K* tiene como (consecuencia lógica)$_{CoP}$ a *X*.

Al lógico Georg Kreisel (véase Kreisel, 1967) se debe un argumento que muestra convincentemente que, en el caso de que las oraciones de *K* y *X* pertenezcan a lenguajes de primer orden, si *X* es una (consecuencia lógica)$_{T}$ de *K*, *K*

tiene como (consecuencia lógica)$_{Cl}$ a X. Kreisel concedió gran importancia a este argumento, pues parece haber identificado la noción intuitiva de consecuencia lógica con la noción generalista de (consecuencia lógica)$_{Cl}$. Pero el argumento puede extenderse fácilmente para mostrar que, también para K y X en primer orden, si X es una (consecuencia lógica)$_T$ de K, K tiene como (consecuencia lógica)$_{ClP}$ a X; y, por tanto (en vista de las implicaciones inmediatas del párrafo precedente), que si X es una (consecuencia lógica)$_T$ de K, entonces K tiene como (consecuencia lógica)$_{Cl}$ a X (lo mostrado originalmente por Kreisel) y además K tiene como (consecuencia lógica)$_{CoP}$ a X. Veamos la extensión del argumento de Kreisel.

La extensión del argumento tiene como premisa la observación de que hay cálculos deductivos para la lógica de primer orden que son *completos* e *intuitivamente correctos* con respecto a la noción de (consecuencia lógica)$_{ClP}$. Los cálculos de Frege y de Russell y Whitehead, de los que hablamos en el capítulo II, son ejemplos de cálculos con estas propiedades (el de Frege con ciertas correcciones menores). Para tener una referencia más accesible, podemos convenir que el cálculo que mencionaremos en nuestra extensión del argumento de Kreisel es el cálculo ofrecido en el capítulo X de Mates (1965). Este cálculo, llamémoslo *Cal*, tiene ocho esquemas formulares, cada uno de los cuales da lugar a un conjunto infinito de axiomas, las clausuras universales de las fórmulas con la misma forma que el esquema. Los esquemas que dan lugar a las fórmulas cuyas clausuras son los axiomas son (véase Mates, 1965: 203):

I. $(\phi\rightarrow(\psi\rightarrow\phi))$;
II. $((\phi\rightarrow(\psi\rightarrow\chi))\rightarrow((\phi\rightarrow\psi)\rightarrow(\phi\rightarrow\chi)))$;
III. $((\neg\psi\rightarrow\neg\phi)\rightarrow(\phi\rightarrow\psi))$;
IV. $(\forall\alpha(\phi\rightarrow\psi)\rightarrow(\forall\alpha\phi\rightarrow\forall\alpha\psi))$;
V. $(\forall\alpha\phi\rightarrow\phi)$;
VI. $(\phi\rightarrow\forall\alpha\phi)$, donde α no aparece libre en ϕ;
VII. $\exists\alpha\alpha=\beta$;
VIII. $(\beta=\gamma\rightarrow(\phi\rightarrow\psi))$, donde ϕ, ψ son atómicas y ψ se diferencia de ϕ solamente en contener una aparición de γ donde ϕ contiene una aparición de β

(donde ϕ, ψ, χ son fórmulas cualesquiera, α es una variable cualquiera y β y γ son símbolos individuales cualesquiera). Como reglas de inferencia, *Cal* tiene

modus ponens e intercambio de signos lógicos por sus definiciones contextuales (por ejemplo, de $\ulcorner\exists x_n\urcorner$ por $\ulcorner\neg\forall x_n\neg\urcorner$). Que *Cal* es completo significa que para cualquier conjunto de oraciones K y cualquier oración X, si K tiene como (consecuencia lógica)$_T$ a X, entonces X es derivable a partir de las oraciones de K en el cálculo *Cal*. La compleción de la lógica de primer orden fue demostrada por primera vez por Gödel (aunque no para el cálculo *Cal*), y constituye un teorema profundo aunque básico de la lógica matemática.

La segunda parte de la premisa de nuestro argumento es que *Cal* es intuitivamente correcto con respecto a la noción de (consecuencia lógica)$_{ClP}$. Esto significa que siempre que X es derivable de K en *Cal*, K tiene como (consecuencia lógica)$_{ClP}$ a X. Para mostrar que esto es así basta observar dos cosas: (1) todos los axiomas de *Cal* son (consecuencias lógicas)$_{ClP}$ de todo conjunto de premisas, o, en otras palabras, son verdaderos en todo (modo de interpretar)$_{ClP}$; y (2) al aplicar una regla de inferencia de *Cal* a una oración o conjunto de oraciones, la oración resultante es (consecuencia lógica)$_{ClP}$ de la oración u oraciones a las que se aplicó. La verificación de que *Cal* es correcto en este sentido sería bastante larga y tediosa, y dejaremos al lector la comprobación de la mayoría de los casos. Aquí sólo verificaremos (1) para el esquema formular IV.

Supongamos, en busca de un absurdo, que hubiese una clausura universal C de una fórmula de la forma IV, $\forall\alpha_1...\forall\alpha_n(\forall\alpha(\phi\rightarrow\psi)\rightarrow(\forall\alpha\phi\rightarrow\forall\alpha\psi))$, que fuera falsa en algún (modo de interpretar)$_{ClP}$. Esto quiere decir que habría un mundo posible m, una clase no vacía U de objetos de m, asignaciones apropiadas de objetos sacados de U a las constantes no lógicas de C y objetos singulares $o_1,...,o_n$ en U tales que, cuando $o_1,...,o_n$ se asignan a $\alpha_1...\alpha_n$ y las constantes no lógicas de C se interpretan por medio de los objetos sacados de U, $(\forall\alpha(\phi\rightarrow\psi)\rightarrow(\forall\alpha\phi\rightarrow\forall\alpha\psi))$ es falsa; o lo que es lo mismo, $\forall\alpha(\phi\rightarrow\psi)$ y $\forall\alpha\phi$ son verdaderas y $\forall\alpha\psi$ es falsa. Pero si $\forall\alpha\psi$ es falsa, hay un objeto o en U tal que cuando o se asigna a α, ψ es falsa. Ahora bien, dado que $\forall\alpha(\phi\rightarrow\psi)$ es verdadera, $(\phi\rightarrow\psi)$ es verdadera cuando o se asigna a α, y dado que $\forall\alpha\phi$ es verdadera, ϕ es verdadera cuando o se asigna a α. Por tanto, ψ ha de ser verdadera cuando o se asigna a α, lo cual contradice lo que ya teníamos. Usando razonamientos similares probaríamos de forma satisfactoria la totalidad de las aserciones que necesitaríamos para establecer (1) y (2). Con ello quedaría establecido que *Cal* es intuitivamente correcto con respecto a la noción de (consecuencia lógica)$_{ClP}$.

Podemos ya ofrecer nuestra extensión del argumento de Kreisel. Tenemos que mostrar que para K y X en primer orden, si X es una (consecuencia lógica)$_T$ de K, K tiene como (consecuencia lógica)$_{ClP}$ a X. Supongamos que X es una (consecuencia lógica)$_T$ de K. Por el teorema de compleción para la lógica de primer orden, esto implica que X es derivable de las oraciones de K por medio del cálculo *Cal*. Pero *Cal* es intuitivamente correcto con respecto a la noción de (consecuencia lógica)$_{ClP}$. Por tanto, K tiene como (consecuencia lógica)$_{ClP}$ a X.

La importancia de este argumento difícilmente puede exagerarse. Muestra que la noción tarskiana de (consecuencia lógica)$_T$, cuando se aplica a lenguajes de primer orden, incorpora el componente de modalidad característico de la noción intuitiva de consecuencia lógica, *al menos* si estamos dispuestos a aceptar alguna de las explicaciones generalistas de la noción intuitiva de consecuencia lógica expuestas en este capítulo. De nuevo vemos que no hay por qué ceder al escepticismo radical sobre la posibilidad del razonamiento intuitivo mixto, con nociones modales como la de (consecuencia lógica)$_{ClP}$ y nociones menos controvertidas como la de (consecuencia lógica)$_T$. Es posible mostrar, como hemos visto, que esta última noción es correcta respecto a la noción modal intuitiva de (consecuencia lógica)$_{ClP}$ para lenguajes de primer orden.

Las limitaciones de este resultado son, sin embargo, obvias. En primer lugar, es un resultado para conjuntos de premisas y conclusiones *de primer orden*. El argumento usa el teorema de compleción para la lógica de primer orden. No hay cálculos efectivos completos con respecto a la noción de (consecuencia lógica)$_T$ para lenguajes de orden superior (y que sean al mismo tiempo correctos con respecto a esa noción). Por tanto, el argumento no se puede reproducir para estos lenguajes. Algunos filósofos de la lógica han negado que todo argumento de segundo orden que sea un caso de (consecuencia lógica)$_T$ sea también un caso de (consecuencia lógica)$_{Cl}$ (y, consiguientemente, que todo argumento que sea un caso de (consecuencia lógica)$_T$ sea también un caso de (consecuencia lógica)$_{ClP}$). Una discusión crítica, que concluye que no se ha demostrado que esos supuestos contraejemplos lo sean y que probablemente no es posible demostrarlo, puede hallarse en Gómez Torrente (1998/9). En el próximo capítulo, que presupone conocimientos relativamente avanzados, daremos algunas razones por las que es plausible pensar que todo argumento de segundo orden que es un caso de (consecuencia lógica)$_T$ es también un caso de (consecuencia lógica)$_{Cl}$.

En segundo lugar, el interés del resultado depende en gran medida de que la noción intuitiva de consecuencia lógica sea analizable por medio de una de las nociones generalistas de consecuencia lógica expuestas en este capítulo. Si no lo es –si, por ejemplo, tenemos razones para pensar que la (consecuencia lógica)$_{ClP}$ no es una condición suficiente de la consecuencia lógica a secas–, entonces tendremos razones para pensar que no hemos mostrado (ni siquiera para lenguajes de primer orden) que la noción de (consecuencia lógica)$_T$ es correcta respecto a la noción intuitiva de consecuencia lógica. En el capítulo IX examinaremos dos nuevas aproximaciones al componente modal de la noción intuitiva de consecuencia lógica, diferentes a las explicadas en este capítulo, y que podrían aducirse para negar que la (consecuencia lógica)$_{ClP}$ sea una condición suficiente de la consecuencia lógica intuitiva. La conclusión, sin embargo, será que en ambos casos la situación resultante es muy similar a la que se plantea respecto a las nociones generalistas de consecuencia lógica discutidas en este capítulo.

VIII. Nociones generalistas de modalidad y lenguajes de segundo orden*

Como adelantamos, ofreceremos aquí algunas razones para pensar que todo argumento de segundo orden que es un caso de (consecuencia lógica)$_T$ es también un caso de (consecuencia lógica)$_{Cl}$. Aunque nos concentraremos en defender esta tesis, al hacerlo no nos estamos olvidando de las tesis análogas de que todo argumento de segundo orden que es un caso de (consecuencia lógica)$_T$ es también un caso de (consecuencia lógica)$_{CoP}$ y de que todo argumento de segundo orden que es un caso de (consecuencia lógica)$_T$ es también un caso de (consecuencia lógica)$_{ClP}$. La razón es que es posible demostrar que todo caso de (consecuencia lógica)$_{Cl}$ es también un caso de (consecuencia lógica)$_{ClP}$ (y, por tanto, de (consecuencia lógica)$_{CoP}$). Y esto es lo que haremos para empezar.

La idea clave de la demostración está contenida en la siguiente proposición:

> *(1) Para todo (modo de interpretar)$_{Cl}$ existe un (modo de interpretar)$_{Cl}$ isomorfo cuyo universo es una clase pura.*

(Una clase pura es una que sólo contiene conjuntos puros –los formados por las operaciones usuales de formación de conjuntos a partir del conjunto

*Este capítulo trata temas algo más avanzados que los introductorios; puede omitirse en una primera lectura.

lógicas de Z por variables nuevas y formar la clausura universal de la fórmula que resulta, obteniendo así una fórmula universal que llamaremos Z'. Z' pertenece al lenguaje de '$\in$' de segundo orden y es (consecuencia lógica)$_T$ de todo conjunto de premisas en ese lenguaje, finito o no; además, Z' ha de ser falsa en I. Pero, por (2), Z' ha de ser verdadera en la clase de los conjuntos puros (pues Z' es (consecuencia lógica)$_T$ de, por ejemplo, el conjunto vacío de premisas). Puesto que Z' no contiene vocabulario no lógico, será falsa en cualquier (modo de interpretar)$_{Cl}$ cuyo universo esté en correspondencia biyectiva con el dominio de I. Pero, como señalamos anteriormente en nuestro argumento para demostrar (1), bajo supuestos razonables toda clase propia está en correspondencia biyectiva con la clase de todos los ordinales. De esto se sigue que toda clase propia está en correspondencia biyectiva con la clase de todos los conjuntos. Así, pues, el dominio de I está en correspondencia biyectiva con el universo de todos los conjuntos y el universo de todos los conjuntos puros. Y entonces Z' ha de ser falsa en la clase de todos los conjuntos puros, lo cual es una contradicción.

(Si aceptáramos la siguiente extensión de (2), que, como veremos en el último párrafo de este capítulo, les ha parecido razonable a varios lógicos eminentes, podríamos dar un argumento para la tesis general de que, para cualesquiera K y X de un lenguaje usual numerable de segundo orden, si X es (consecuencia lógica)$_T$ de K, entonces X es (consecuencia lógica)$_{Cl}$ de K. La extensión de (2) es

> *(3) Todos los argumentos que son casos de (consecuencia lógica)$_T$ en el lenguaje infinitario de cardinalidad $\aleph_0$ de segundo orden con '$\in$' son tales que o alguna premisa es falsa o la conclusión es verdadera en la clase de los conjuntos puros (con la relación de pertenencia).*

El argumento es el siguiente. Supongamos que X es una (consecuencia lógica)$_T$ de K pero también que hay un (modo de interpretar)$_{Cl}$ I tal que todas las oraciones de K son verdaderas bajo I pero X es falsa bajo I. Formemos un condicional Z con una conjunción (posiblemente con $\aleph_0$ miembros) de las oraciones de K como antecedente y X como consecuente. Sustituyamos de manera uniforme todas las constantes no lógicas de Z por variables nuevas y formemos la clausura universal, Z'. Como antes, Z' ha de ser falsa en I. Y por

(3), Z' es verdadera en la clase de los conjuntos puros (pues Z' es una (verdad lógica)$_T$). Pero, además, por un argumento análogo al de antes, Z' ha de ser falsa en la clase de todos los conjuntos puros, que es el absurdo buscado.)

En vista de estas consideraciones, podemos concentrarnos de manera justificada en (2). Un argumento de segundo orden que podría a primera vista considerarse como un posible contraejemplo a (2) (y, por tanto, a (3)) sería uno que tuviera como premisas el conjunto finito de axiomas de segundo orden de Zermelo-Fraenkel (ZF) y como conclusión una oración T cualquiera que fuera claramente falsa en la interpretación natural de la teoría de conjuntos, por ejemplo, '$\exists x \neg x=x$'. Pues sabemos que un (modo de interpretar)$_{Cl}$ del lenguaje de '$\in$' hace verdaderos los axiomas de segundo orden de Zermelo-Fraenkel si y sólo si es isomorfo a la clase de todos los conjuntos (puros) de rango menor que el de un cardinal inaccesible, con la relación de pertenencia (a veces se usa el término "cuasicategórica" para referirse a la propiedad de la teoría de Zermelo-Fraenkel de segundo orden de caracterizar a la colección de todas las interpretaciones de un cierto tipo). Por tanto, si no hay cardinales inaccesibles, el argumento

$$\begin{array}{l} Ax_1 \\ Ax_2 \\ \textit{etc.} \\ \hline T, \end{array}$$

(donde Ax_1, Ax_2, etc., son los axiomas de ZF de segundo orden) es un caso de (consecuencia lógica)$_T$ pero no un caso de (consecuencia lógica)$_{Cl}$ (pues ZF sería verdadera, y T falsa, en el (modo de interpretar)$_{Cl}$ proporcionado por la clase de todos los conjuntos de rango inferior al menor cardinal inaccesible). Sin embargo, la mayoría de los matemáticos que trabajan en teoría de conjuntos creen que los cardinales inaccesibles son conjuntos y, por tanto, que el argumento en cuestión no es un caso de (consecuencia lógica)$_T$.

Hay otros argumentos de segundo orden tales que las oraciones del conjunto (finito) formado por las premisas son conjuntamente verdaderas en y sólo en (o, en otras palabras, cuasicaracterizan) ciertas colecciones de interpretaciones de la forma V_α (V_α es el modo de interpretar la teoría de conjuntos que

consta de los conjuntos de rango menor que el ordinal α). Pero la actitud general entre los que trabajan en teoría de conjuntos parece consistir en seguir un principio informal de maximización del contenido del universo de la teoría de conjuntos, principio que lleva a aceptar todas esas grandes interpretaciones como conjuntos.

Hay un teorema de la teoría estándar de conjuntos que garantiza que la potencia del lenguaje de '$\in$' de segundo orden para caracterizar (y cuasicaracterizar) grandes (modos de interpretar)$_{Cl}$ es limitado. El teorema dice que si la colección de fórmulas de un lenguaje formal L de la jerarquía clásica de lenguajes de orden finito es un conjunto, entonces hay un menor cardinal κ tal que, para todo conjunto de oraciones K de L, si todas las oraciones de K son verdaderas en una estructura para L, entonces son también verdaderas en una estructura cuyo universo tiene cardinalidad menor o igual que κ (el primero en reparar en esto fue probablemente William Hanf). La prueba es relativamente simple (pero da escasa información sobre el tipo de cardinal que κ pueda ser en cada caso). Sea Σ la colección de los conjuntos de oraciones de L que son conjuntamente verdaderas en alguna estructura. Si C es un conjunto en Σ, sea κ_C el menor cardinal tal que todas las oraciones de C son verdaderas en una estructura con universo de cardinalidad κ_C. Puesto que la colección de las oraciones de L es un conjunto, su conjunto potencia y por tanto Σ son conjuntos. Luego los κ_C con C en Σ forman un conjunto, por el axioma de reemplazo. La unión o supremo de este conjunto es un cardinal κ con las propiedades deseadas. Este cardinal recibe usualmente el nombre de "número de Löwenheim para L", porque κ desempeña con respecto a L un papel análogo al que $\aleph_0$ desempeña con respecto a los lenguajes de primer orden en el teorema de Löwenheim-Skolem.

Claramente, el número de Löwenheim del lenguaje de '$\in$' de segundo orden ha de ser muy grande (al menos si adoptamos la actitud "maximizante" que prevalece entre los principales teóricos de conjuntos), en vista de su potencia para cuasicaracterizar colecciones de grandes (modos de interpretar)$_{Cl}$ del tipo V_α. Aunque poco se sabe acerca de las propiedades intrínsecas de este número de Löwenheim, se sabe que es mucho más pequeño que el menor de ciertos grandes cardinales considerados por los matemáticos, los llamados cardinales supercompactos. De nuevo la tendencia preponderante entre los teóricos de conjuntos es incluir a los cardinales supercompactos entre los conjuntos.

El teorema de que el lenguaje de '$\in$' de segundo orden tiene un número de Löwenheim garantiza que hay un cierto rango ordinal ρ tal que, si una oración X es verdadera en todas las estructuras de rango menor que ρ en las que todas las oraciones de un conjunto K son verdaderas, entonces X es verdadera en todas las estructuras en que todas las oraciones de K son verdaderas, a secas (es decir, si X es verdadera en todas las estructuras de rango menor que ρ en las que todas las oraciones de K son verdaderas, entonces X es (consecuencia lógica)$_T$ de K). Por tanto, garantiza que sólo una parte relativamente pequeña (aunque muy grande en términos absolutos) de los pisos bajos del universo conjuntista es necesaria para proporcionar estructuras que refuten los argumentos que no son casos de (consecuencia lógica)$_T$. Esto podría interpretarse como evidencia de la riqueza del universo de la teoría de conjuntos e, indirectamente, de la probabilidad de (2): el proceso de formación de conjuntos es tan pródigo que después de un cierto punto ha producido contraejemplos para todos los argumentos que no son casos de (consecuencia lógica)$_T$; es improbable, por tanto (o así continuaría esta línea de razonamiento), que extender el proceso a un proceso de formación de clases produjera (modos de interpretar)$_{Cl}$ que refutaran algunos nuevos argumentos no refutados por ninguna estructura conjuntista. Desde luego, este argumento está lejos de ser concluyente, aunque no carece por entero de fuerza.

(Es preciso prevenirse contra cierto argumento incorrecto con la misma conclusión. Es análogo a un argumento incorrecto acerca de los lenguajes de primer orden que parece haber sido defendido por algunos en la bibliografía filosófica. Con el argumento incorrecto acerca de los lenguajes de primer orden se pretende establecer que las verdades de la teoría de conjuntos de primer orden son verdaderas en alguna interpretación numerable, aplicando el teorema de Löwenheim-Skolem. Este teorema, en su forma pertinente usual, dice que todo conjunto de oraciones de primer orden verdaderas en alguna estructura (conjuntista, por tanto) también es tal que todas sus oraciones son verdaderas en alguna estructura con universo numerable. Si bien es una premisa auxiliar aceptable que las verdades de la teoría de conjuntos de primer orden son verdaderas en algún (modo de interpretar)$_{Cl}$, uno no puede concluir de esto y del teorema de Löwenheim-Skolem, tal como se lo formula usualmente, que la teoría de conjuntos es verdadera en una teoría numerable. Un argumento correcto para defender esta afirmación usa el teorema de compleción de Gödel, que garantiza

que un conjunto consistente de oraciones de primer orden (por ejemplo, consistente según el cálculo del capítulo VII) es verdadero en una estructura numerable, junto con la plausible premisa auxiliar de que la teoría de conjuntos es consistente. Puesto que verdad en alguna interpretación implica consistencia –el teorema de corrección para un cálculo–, el teorema de Löwenheim-Skolem se demuestra a menudo usando el teorema de compleción de Gödel. Pero el *enunciado* del teorema de Löwenheim-Skolem sólo puede ser usado en el argumento de más arriba si uno está dispuesto a cometer un error. Si pasáramos por encima de este error sin notarlo, podríamos sentirnos tentados a usar un argumento similar en el caso de los lenguajes de segundo orden, pues, como hemos indicado, hay un teorema para estos lenguajes análogo al teorema de Löwenheim-Skolem. Este argumento incorrecto concluiría que los argumentos que no son casos de $(\text{consecuencia lógica})_{Cl}$ de la teoría de conjuntos de segundo orden han de ser refutables en una estructura cuya cardinalidad fuera el número de Löwenheim del lenguaje de '$\in$' de segundo orden, partiendo del supuesto de que han de ser refutables en el $(\text{modo de interpretar})_{Cl}$ la teoría de conjuntos que es natural para ésta.)

Además de los anteriores, se han propuesto otros ejemplos de argumentos del lenguaje de '$\in$' de segundo orden de los que se ha defendido que constituyen contraejemplos a la tesis (2) (y, por tanto, a (3)), pero no los discutiremos aquí. (Una propuesta especialmente interesante de Vann McGee (1992) se critica en Gómez Torrente, 1998/9.) Sin embargo, es muy importante señalar, independientemente de consideraciones de detalle, que la idea de que hay conjuntos de oraciones verdaderas en la interpretación natural de la teoría de conjuntos, pero no satisfacibles en ninguna estructura conjuntista, va en contra de otro principio informal de los que siguen los teóricos de conjuntos en la exploración de nuevos axiomas para la teoría de conjuntos. Este principio es llamado a veces "principio de reflexión". Una formulación particularmente potente la da Hao Wang, quien se la atribuye a Gödel:

> "Principio de reflexión: el universo de los conjuntos es estructuralmente indefinible. Una manera posible de hacer precisa esta afirmación es la siguiente: El universo de los conjuntos no puede ser unívocamente caracterizado (esto es, distinguido de todos sus segmentos iniciales) por medio de ninguna propiedad estructural interna de la relación $\in$ sobre ese universo,

> expresable en una lógica de tipo finito o transfinito, incluyendo lógicas infinitarias de cualquier cardinalidad" (Wang, 1974: 540).

La idea es que

> "lo que sea verdadero del universo entero ha de ser verdadero ya de algún segmento inicial del universo. En otras palabras, todo intento de describir unívocamente V [el universo de los conjuntos] también se aplica a [V_α's] más pequeños que "reflejan" la propiedad adscrita a V" (Maddy, 1988: 503).

Versiones del principio de reflexión para clases de fórmulas de segundo orden de complejidad relativamente baja han sido exploradas por teóricos significados, empezando por Bernays, y parecen ser confirmadas por el hecho de que implican muchos teoremas aceptados de la teoría estándar de conjuntos. En conclusión, pues, puede decirse que aunque no hay argumentos concluyentes de ninguno de los dos lados, la carga de la prueba claramente recae en este momento en aquellos que quieren negar la tesis de que todo argumento de segundo orden que es un caso de $(\text{consecuencia lógica})_T$ es también un caso de $(\text{consecuencia lógica})_{Cl}$ negando (2), pues la práctica de los lógicos se basa a menudo en el supuesto de que (2) (y también versiones de (2) mucho más potentes aún, como (3)) es verdadera. Aunque esto no constituye un argumento definitivo en favor de la tesis, la hace más probable que la tesis contraria.

IX. Consecuencia lógica, aprioridad, analiticidad

Cuando un conjunto de premisas K tiene como (consecuencia lógica)$_{Cl}$ a una conclusión X, o simplemente cuando K tiene como (consecuencia lógica)$_{T}$ a X (hemos visto que para lenguajes de primer orden los dos conceptos son coextensionales), ello no parece llevar consigo sin más que X es obtenible a partir de K por razonamiento *a priori*, es decir, que K implica X *a priori*. Sin embargo, cuando X es consecuencia lógica en el sentido intuitivo de K, parece que ha de ser condición necesaria de ello que K implique X *a priori*, y, además, que el razonamiento *a priori* que lleva de K a X incluya únicamente reflexión *a priori* sobre los conceptos denotados por las constantes lógicas. La noción intuitiva de consecuencia lógica y la de implicación por necesidad lógica (en términos de la cual hemos buscado expresar el componente modal del concepto de consecuencia lógica) parecen poder entenderse, en parte, por medio de la noción modal de implicación *a priori*. ¿Hay un conjunto de oraciones K y una oración X tales que K tiene como (consecuencia lógica)$_{T}$ a X pero K no implica X *a priori*? Si fuera así, habría buenos motivos para negar la corrección de la definición tarskiana.

La respuesta más simple a la pregunta es "sí". Hay oraciones de lenguajes de primer orden que son (verdades lógicas)$_{T}$ y, por tanto, (consecuencias lógicas)$_{T}$ de todo conjunto de premisas, pero que no parecen ser implicadas *a priori* por todo conjunto de premisas. Por ejemplo, las oraciones '$\exists x x=x$', '$\exists x(Nx\rightarrow Nx)$' y '$\exists x 0=x$' son (verdades lógicas)$_{T}$, ya que son verdaderas en toda estructura, al

tener siempre una estructura un conjunto *no vacío* como universo, y al convenirse que toda constante individual tiene asignado un objeto al reinterpretar el lenguaje al que pertenece en cada universo no vacío, una *denotación*. Sin embargo, no son verdades *a priori* bajo todos los modos de interpretar sus constantes no lógicas. Es decir, aunque podemos interpretar el recorrido de las variables y las constantes de forma que '$\exists xx=x$' signifique "hay un número natural idéntico a sí mismo" y '$\exists x0=x$' signifique "hay un número natural idéntico a cero", caso en el cual sí parecen ser verdades *a priori*, también podemos interpretarlos de forma que '$\exists xx=x$' signifique "hay un filósofo idéntico a sí mismo" y '$\exists x0=x$' signifique "hay un filósofo idéntico a Aristóteles", y en estos casos las oraciones no parecen verdades *a priori*. (Nótese que si '$\exists xx=x$' significase "hay un objeto singular idéntico a sí mismo" a secas, también parecería *a priori*, pues "hay un número natural idéntico a sí mismo" parece *a priori*, y parece implicar *a priori* "hay un objeto singular idéntico a sí mismo".)

La importancia de estos ejemplos y otros similares no se debe exagerar. El hecho de que sean (verdades lógicas)$_T$ se debe a que la noción de estructura en que se basa (CLT) no incluye "estructuras" con universo vacío. Sin embargo, nada habría impedido que hubiésemos considerado una noción de estructura que abarcase tales "estructuras", y que con su ayuda hubiésemos definido nociones más estrictas de (consecuencia lógica)$_\#$ y (verdad lógica)$_\#$, tales que '$\exists xx=x$' y oraciones similares no serían (verdades lógicas)$_\#$ ni (consecuencias lógicas)$_\#$ de todo conjunto de premisas. Al no dar esas definiciones alternativas seguimos la práctica usual de los lógicos cuando enuncian las definiciones tarskianas. El motivo principal de esa práctica de los lógicos es de orden pragmático. La colección de las estructuras con universo no vacío contiene todas las estructuras interesantes consideradas por los lógicos y matemáticos, y las oraciones verdaderas en todas las estructuras de esa colección forman un conjunto interesante, algo más amplio que el de las oraciones verdaderas en todas las "estructuras", con universo vacío o no vacío. Además, la formulación de cálculos deductivos se simplifica notablemente bajo la suposición de que es permisible introducir en las derivaciones fórmulas verdaderas en toda estructura pero no en toda "estructura".

Esta situación no genera una objeción seria a la corrección de (CLT) o (VLT). Las definiciones alternativas de (consecuencia lógica)$_\#$ y (verdad lógica)$_\#$ a las que hemos aludido podrían darse fácilmente con levísimas modificaciones de las

que ya dimos. De hecho, han sido ofrecidas por varios autores y se han dado cálculos de primer orden correctos y completos con respecto a tales nociones. Pero esas definiciones y esos cálculos no aparecen en los textos usuales de lógica y son poco conocidos. Por ello, es preferible presentar nuestros argumentos a propósito de la noción usual de (consecuencia lógica)$_T$ y de cálculos usuales como *Cal*.

Lo que sí sería un problema serio para la tesis de la corrección de (CLT) o (VLT) sería la existencia de un conjunto de oraciones K y una oración X tales que K tuviera como (consecuencia lógica)$_T$ a X pero K no implicara X *a priori*, ni siquiera *bajo el supuesto adicional* de que hay al menos un objeto en el universo del discurso y toda constante individual tiene denotación. Si esto fuera así, no habría a primera vista ninguna razón para pensar que tal situación se debiera a algún supuesto pragmático y prescindible empleado en la formulación de (CLT). Más bien, tendríamos una buena razón para pensar que (CLT) es incorrecta por motivos fundamentales. Por ello es de gran importancia observar que es de nuevo posible extender el argumento de Kreisel que ya extendimos en el capítulo VII, y convertirlo en un argumento para mostrar que (para K y X de primer orden) si K tiene como (consecuencia lógica)$_T$ a X, entonces K implica X *a priori* bajo el supuesto adicional de que hay al menos un objeto en el universo del discurso y toda constante individual tiene denotación.

Veamos el nuevo argumento. Su primera premisa es que *Cal* es intuitivamente correcto con respecto a la noción de implicación *a priori*, bajo el supuesto adicional de que hay al menos un objeto en el universo y toda constante individual tiene denotación. Es decir: para cualquier conjunto de oraciones K y cualquier oración X (de primer orden), si X es derivable a partir de las oraciones de K en el cálculo *Cal*, entonces K implica X *a priori* bajo el supuesto adicional de que hay al menos un objeto en el universo y toda constante individual tiene denotación. Esto se justifica de nuevo por inspección: las reglas de *Cal* llevan de oraciones a oraciones implicadas *a priori* por las primeras, y los axiomas de *Cal* son todos o verdades *a priori* o son implicados *a priori* por la proposición de que hay al menos un objeto en el universo del discurso y toda constante individual tiene denotación. En particular, los axiomas del grupo VII, las clausuras de fórmulas de la forma $\exists\alpha\alpha=\beta$, no son siempre verdades *a priori*, pero son implicados *a priori* por aquella proposición (si consideramos lenguajes de primer orden con símbolos de función, deberemos enriquecer

esta proposición y añadir la premisa de que toda constante funcional tiene como denotación una función total en el dominio para el que está definida).

Ahora bien, el teorema de compleción para la lógica de primer orden nos dice que *Cal* es completo respecto a la noción de (consecuencia lógica)$_T$: para cualquier conjunto de oraciones K y cualquier oración X (de primer orden), si K tiene como (consecuencia lógica)$_T$ a X, entonces X es derivable a partir de las oraciones de K en el cálculo *Cal*. Esta es la segunda premisa del argumento. Junto con la primera, implica que para cualquier conjunto de oraciones K y cualquier oración X (de primer orden), si K tiene como (consecuencia lógica)$_T$ a X, entonces K implica X *a priori* bajo el supuesto adicional de que hay al menos un objeto en el universo y toda constante individual tiene denotación. Este es un resultado importante, pues muestra que, para el caso de lenguajes de primer orden, la noción de (consecuencia lógica)$_T$ es correcta respecto a la noción de implicación *a priori*, salvo en casos incluidos por razones pragmáticas y fácilmente prescindibles.

Pero de nuevo resultan claras las también importantes limitaciones de esta observación. Es una observación restringida a conjuntos de premisas y conclusiones de primer orden y, dada la incompleción de todo cálculo para la noción de (consecuencia lógica)$_T$ en lenguajes de orden superior, el argumento no se puede reproducir para estos lenguajes. Con cierta frecuencia se encuentran argumentos que niegan que toda oración de segundo orden que es (consecuencia lógica)$_T$ de todo conjunto de premisas es implicada *a priori* por todo conjunto de premisas (incluso bajo el supuesto adicional de que hay al menos un objeto en el universo y toda constante individual tiene denotación). La discusión crítica de algunos de estos argumentos nos llevaría fuera de los límites de esta monografía, pero el núcleo central de uno de ellos puede ser criticado aquí sin salirnos demasiado de nuestro camino.

El argumento que queremos criticar podría presentarse aproximadamente como sigue. Todo cálculo para la noción de (consecuencia lógica)$_T$ en lenguajes de orden superior es incompleto. Es decir, no hay ningún conjunto efectivamente generable de axiomas y reglas efectivas que produzca mecánicamente todas las (consecuencias lógicas)$_T$ de conjuntos de premisas en lenguajes de orden superior. Además, no hay ningún cálculo que produzca mecánicamente todas las (consecuencias lógicas)$_T$ del conjunto vacío de premisas, es decir, todas las (verdades lógicas)$_T$. Pero el razonamiento *a priori* ha de ser reproducible

de forma efectiva, si es que ha de ser posible que lo efectúen seres que razonen. Por tanto, concluye el argumento, ha de haber casos de (verdad lógica)$_T$ que no son casos de verdad *a priori* (ni siquiera bajo el supuesto adicional de que hay al menos un objeto en el universo y toda constante individual tiene denotación).

Este razonamiento, a pesar de lo frecuente que es, es claramente incorrecto. Supongamos que aceptamos que (1) el razonamiento *a priori* ha de ser efectuable por medio de un cálculo. Aceptamos también, naturalmente, que (2) *para todo* cálculo *existen* (verdades lógicas)$_T$ que no son derivables sin premisas en él. Pero de esto no se sigue que (3) *existen* (verdades lógicas)$_T$ tales que *para todo* cálculo, no son derivables sin premisas en ese cálculo. De (3) y (1) se sigue, desde luego, que hay (verdades lógicas)$_T$ que no son verdades *a priori*, y, por tanto, casos de (consecuencia lógica)$_T$ que no son casos de implicación *a priori*. Pero el paso de (2) a (3) es incorrecto (se basa en un error típico producido por la ambigüedad de ciertos enunciados con varios cuantificadores en el lenguaje natural, que se aclara en la formalización distinguiendo entre maneras de ordenar los cuantificadores). De (1) y (2) no se sigue que exista ninguna proposición que sea una (verdad lógica)$_T$ pero no sea cognoscible *a priori*; y, sin embargo, algunos proponentes del argumento han dicho que se sigue.

Como en nuestra primera extensión del argumento de Kreisel, el interés de la nueva extensión que hemos hecho en este capítulo será mayor o menor dependiendo de la relación exacta entre las nociones de implicación *a priori* y de consecuencia lógica intuitiva. ¿Hay alguna noción más estricta si cabe que la noción de implicación *a priori* y que suministre una condición necesaria de todo argumento que sea un caso intuitivo de consecuencia lógica? En lo que resta de este capítulo consideraremos una noción que a primera vista parece tener tales propiedades, la noción que llamaremos de *implicación analítica*.

Digamos que un conjunto de premisas K implica analíticamente una conclusión X cuando K implica X *en virtud del significado de las constantes lógicas* del lenguaje al que pertenecen K y X. (Este uso de "implicación analítica" no es estándar; normalmente se habla de implicación analítica para referirse a la implicación en virtud del significado de las expresiones de un argumento, sean constantes lógicas o no. Las implicaciones analíticas en nuestro sentido son un subconjunto propio de las implicaciones analíticas en el sentido estándar. Usamos el término por comodidad.) ¿Hay K y X tales que K tiene como (consecuencia lógica)$_T$ a X pero K no implica

analíticamente X? Como respuesta se pueden hacer consideraciones enteramente análogas a las que hemos hecho anteriormente en este capítulo a propósito de la relación entre (consecuencia lógica)$_T$ e implicación *a priori*. Hay oraciones, como '$\exists x x=x$' y '$\exists x 0=x$', que son (verdades lógicas)$_T$ pero no son analíticas (entendemos aquí por "analíticas" "verdaderas en virtud del significado *de sus constantes lógicas*") bajo todos los modos de interpretarlas. Pero de nuevo puede mostrarse que la objeción no es de gran importancia, pues, usando una nueva extensión del argumento de Kreisel, enteramente análoga a la anterior, es posible persuadirnos de que para K y X de primer orden, si K tiene como (consecuencia lógica)$_T$ a X entonces K implica analíticamente X bajo el supuesto adicional de que hay al menos un objeto en el universo del discurso y toda constante individual tiene denotación. (Un argumento esencialmente idéntico, pero que no hace la importante salvedad revelada por el supuesto adicional, y es, por tanto, deficiente, puede hallarse en Etchemendy, 1990: 147.) La comprobación de que esta justificación es posible se deja al lector.

¿Hay K y X *en lenguajes de orden superior* tales que K tiene como (consecuencia lógica)$_T$ a X pero K no implica analíticamente X? Esta es una cuestión que no podemos tratar con la profundidad necesaria dentro de los confines de este trabajo. Pero sí es posible exponer algo superficialmente un argumento reciente que concluye que hay (verdades lógicas)$_T$ de lenguajes de orden superior que no son analíticas, e indicar por qué ese argumento es incorrecto. El argumento ha sido propuesto por John Etchemendy (véase Etchemendy, 1990: 123-124 y 132).

La primera premisa del argumento es que hay dos oraciones de *segundo orden* A y B tales que

(a) A es una (verdad lógica)$_T$ si y sólo si HC;
(b) B es una (verdad lógica)$_T$ si y sólo si $\neg$HC.

Aquí 'HC' está en lugar de una formulación natural de una proposición algo complicada enunciable en el lenguaje de la teoría de conjuntos, la "hipótesis del continuo", que dice que no hay conjuntos de cardinalidad intermedia entre la del conjunto de los números naturales y la del conjunto de los conjuntos de números naturales. Las equivalencias (a) y (b) son en efecto verdaderas y

demostrables en la teoría de conjuntos. El hecho de que HC y su negación son complicadas lleva a nuevas premisas del argumento de Etchemendy:

(c) la hipótesis del continuo no es conceptualmente verdadera;
(d) la negación de la hipótesis del continuo no es conceptualmente verdadera.

(donde "conceptualmente verdadera" significa "verdadera en virtud del significado de sus expresiones, tanto lógicas como no lógicas"). Por último, la premisa final es

(e) para toda oración O, si no es conceptualmente verdadero que [O es una (verdad lógica)$_T$] entonces O no es analítica.

De (a) y (c) Etchemendy concluye

(f) no es conceptualmente verdadero que [A es una (verdad lógica)$_T$];

y de (b) y (d) concluye

(g) no es conceptualmente verdadero que [B es una (verdad lógica)$_T$].

De (e), (f) y (g) se sigue que ni A ni B son analíticas. Sin embargo, o A es una (verdad lógica)$_T$ o B es una (verdad lógica)$_T$, ya que o HC es verdadera o $\neg$HC lo es (por lógica básica). Así, pues, concluye el argumento, hay (verdades lógicas)$_T$ que no son analíticas.

El argumento se basa en una premisa extremadamente dudosa, (e); una discusión crítica, que sugiere que (e) es falsa y que Etchemendy –que no la justifica– la introduce por una simple confusión, puede hallarse en Gómez Torrente (1998/9). Sin embargo, el argumento es incorrecto incluso si no ponemos en duda (e). El error es que (f) y (g) *no se siguen* de (a), (b), (c) y (d). Para concluir de (i) "p si y sólo si q" y (ii) "q no es conceptualmente verdadero" que (iii) "p no es conceptualmente verdadero", la equivalencia entre p y q afirmada por (i) ha de ser una equivalencia conceptual, no una equivalencia puramente material o a lo sumo matemáticamente necesaria, como las que se afirman en (a) y (b). Quizá un ejemplo con otra modalidad ayude a percibir mejor el error. La oración

2+2=4 si y sólo si doña Pura pesa treinta kilos de más

es verdadera: es un bicondicional material cuyos dos miembros son verdaderos. Además,

'Doña Pura pesa treinta kilos de más' no es metafísicamente necesario

es presumiblemente verdadero. Pero de esas dos oraciones no se sigue la oración

'2+2=4' no es metafísicamente necesario,

que es presumiblemente falsa. (Aunque se seguiría, naturalmente, de

'Doña Pura pesa treinta kilos de más' no es metafísicamente necesario

y

'2+2=4 si y sólo si doña Pura pesa treinta kilos de más' es metafísicamente necesario,

lo cual, por suerte para doña Pura, no es verdadero.)

Concluimos aquí nuestra discusión de la segunda parte del problema de la corrección de (CLT). Como vimos en el capítulo V la noción de (consecuencia lógica)$_T$ (o la noción coextensional de (consecuencia lógica)$_t$) es, sin duda, formal para los lenguajes de primer orden, y también para lenguajes de órdenes superiores. En los capítulos VI, VII y IX hemos visto que hay razones convincentes para pensar que, para lenguajes de primer orden, es también una noción modal, desde distintas formas de entender el componente modal del concepto intuitivo de consecuencia lógica; y, además, en los capítulos VIII y IX hemos indicado que no hay buenos motivos para pensar que en lenguajes de orden superior deje de ser modal en ninguno de esos distintos sentidos. Todo hace pensar que la búsqueda de motivos mejores que los ofrecidos hasta ahora (y discutidos con cierto detalle en Gómez Torrente, 1998/9) ha de ser una empresa difícil, si no imposible.

X. EL CONCEPTO DE CONSTANTE LÓGICA Y SUS DIFICULTADES

Al definir la noción de (consecuencia lógica)$_T$ para un lenguaje, presuponemos que se ha fijado de antemano un cierto conjunto de sus constantes como el conjunto de sus *constantes lógicas* (o, indiferentemente, que se ha fijado el complemento de ese conjunto, el conjunto de las constantes *no* lógicas). La relación de (consecuencia lógica)$_T$ variará según varíe el conjunto aceptado de constantes lógicas. Si aceptáramos que el conjunto de constantes lógicas de un lenguaje es uno diferente del que ordinariamente aceptamos, la relación definida de (consecuencia lógica)$_T$ será también diferente a la generada por la definición basada en el conjunto de constantes lógicas ordinariamente aceptado.

Veamos un ejemplo con detalle. Consideremos un lenguaje de primer orden con las mismas constantes lógicas que *LAr*, es decir, '$\forall$', '$\rightarrow$', '$\neg$' y '=', las mismas variables y signos auxiliares, '(', ')', 'x', y ',', y que como constantes no lógicas tiene a la constante individual 'π', el predicado monádico 'Un', y el predicado diádico 'Fu'. Este lenguaje –llamémoslo *LUn*– tiene la misma composición estructural que *LAr*, si bien el significado de sus constantes no lógicas es diferente: 'π' denota en la interpretación deseada al número pi, 'Un' significa "es un unicornio", y 'Fu' significa "fusila a". Al ser estructuralmente idéntico a *LAr*, las definiciones de (consecuencia lógica)$_T$ y (verdad lógica)$_T$ para *LUn* se darán de forma enteramente paralela a las correspondientes definiciones para *LAr*. El resultado será que un argumento de *K* a *X* en *LAr* es un caso de (consecuencia lógica)$_T$ si y sólo si el argumento de *K'* a *X'* en *LUn* es un caso

de (consecuencia lógica)$_T$ (donde K' y X' se obtienen de K y X sustituyendo uniformemente '0' por 'π', 'N' por 'Un' y 'M' por 'Fu').

Pero entonces supongamos que aceptáramos que el conjunto de las constantes lógicas de *LUn* no es el que acabamos de decir y ordinariamente aceptaríamos, sino {'$\forall$', '$\neg$', 'Un', 'Fu'}. En ese caso el conjunto de constantes no lógicas será {'π', '=', '$\rightarrow$'}. Una *estructura* apropiada para *LUn* será entonces una secuencia que tiene como primer miembro a un conjunto no vacío U y que asigna objetos "sacados" de U a las constantes no lógicas de *LUn*; pero los objetos apropiados serían, en este caso: un miembro a de U asignado a 'π', una relación binaria S entre objetos de U asignada a '=' y, asignada a '$\rightarrow$', una operación G definida sobre el conjunto de pares ordenados de conjuntos de secuencias (de las que asignan valores en U a las variables de *LUn*), y que escoge para cada par un conjunto de secuencias (también que asignen valores en U a las variables de *LUn*). (La idea es que operaciones como G son los posibles valores semánticos de una conectiva, entre ellos, funciones de verdad. Un ejemplo: llamemos 'U^ω' al conjunto de las secuencias que asignan valores en U a las variables de *LAr*; el condicional material sería la interpretación de '$\rightarrow$' consistente en la operación G que asigna a un par ordenado $<S_1, S_2>$ de subconjuntos de U^ω el subconjunto $(U^\omega\text{-}S_1)\cup S_2$ de U^ω –recuérdese que $\ulcorner(P\rightarrow Q)\urcorner$ puede definirse por medio de $\ulcorner(\neg P\vee Q)\urcorner$.)

Una *función formular F'* de una fórmula F de *LUn* será ahora el resultado de sustituir 'π' por la (nueva) variable 'y', '=' por la (nueva) variable 'I' y '$\rightarrow$' por la (nueva) variable 'c' de una manera uniforme en F. El concepto de satisfacción de una función formular por una estructura se definirá entonces así: *la estructura* $<U, a, S, G>$ *satisface la función formular* X *con respecto a una secuencia* f (que asigne valores en U a las variables de *LUn*) si y sólo si:

> *(α) (i)* X *es* $\ulcorner Unx_n\urcorner$ *(para algún* n*) y* $f(x_n)$ *es un unicornio; o* X *es 'Uny' y* a *es un unicornio; o*
> *(ii)* X *es* $\ulcorner Fux_nx_m\urcorner$ *(para algunos* m *y* n*) y* $f(x_n)$ *fusila a* $f(x_m)$*; o* X *es* $\ulcorner Fuyx_n\urcorner$ *(para algún* n*) y* a *fusila a* $f(x_n)$*; o* X *es* $\ulcorner Fux_ny\urcorner$ *(para algún* n*) y* $f(x_n)$ *fusila a* a*; o* X *es 'Fuyy' y* a *fusila a* a*; o*
> *(iii)* X *es* $\ulcorner Ix_nx_m\urcorner$ *(para algunos* m *y* n*) y* $<f(x_n), f(x_m)>\in S$*; o* X *es* $\ulcorner Iyx_n\urcorner$ *(para algún* n*) y* $<a, f(x_n)>\in S$*; o* X *es* $\ulcorner Ix_ny\urcorner$ *(para algún* n*) y* $<f(x_n), a>\in S$*; o* X *es 'Iyy' y* $<a, a>\in S$*; o*

> *(β) hay una función formular* Y *tal que* X *es* $\ulcorner\neg Y\urcorner$ *y* $<U, a, S, G>$ *no satisface* Y *con respecto a la secuencia* f*; o*
> *(γ) hay funciones formulares* Y *y* Z *tales que* X *es* $\ulcorner(Y\subset Z)\urcorner$*,* S_1 *es el conjunto de secuencias* $\{g: g\in U^\omega$ *y* $<U, a, S, G>$ *satisface* Y *con respecto a la secuencia* $g\}$*,* S_2 *es el conjunto de secuencias* $\{g: g\in U^\omega$ *y* $<U, a, S, G>$ *satisface* Z *con respecto a la secuencia* $g\}$*, y* $f\in G(<S_1, S_2>)$*; o, por último,*
> *(δ) hay una función formular* Y *y un número* n *tales que* X *es* $\ulcorner\forall x_nY\urcorner$ *y toda secuencia* g *que asigna valores en* U *a las variables de LAr y que difiere de* f *a lo sumo en lo que asigna a* x_n *es tal que* $<U, a, S, G>$ *satisface* Y *con respecto a* g*.*

Se trata nuevamente de una definición recursiva como las otras que ya vimos y en la que se han hecho los cambios necesarios. A partir de esta definición daríamos definiciones exactamente iguales a las del capítulo VI para las nociones de satisfacción de una función oracional, estructura modelo, y, por último, (consecuencia lógica)$_T$ y (verdad lógica)$_T$.

Bajo la selección de constantes lógicas en que se basan estas definiciones, ciertos argumentos que eran casos de (consecuencia lógica)$_T$ bajo la selección ordinaria no son ahora casos de (consecuencia lógica)$_T$; y viceversa, ciertos argumentos que no son casos de (consecuencia lógica)$_T$ bajo la selección ordinaria son ahora casos de (consecuencia lógica)$_T$. Por ejemplo, el argumento

$$\frac{(\forall x\neg Unx\rightarrow\exists x\exists x Fuxx)\quad \forall\neg Unx}{\exists x\exists x Fuxx.}$$

(que es correcto por *modus ponens*) no es un caso de (consecuencia lógica)$_T$ bajo la nueva selección (para verlo, tómese por ejemplo una estructura cuyo universo U es un conjunto de números naturales y donde '$\rightarrow$' denota la operación G definida sobre $U^\omega\times U^\omega$ que asigna a un par $<S_1, S_2>$ el conjunto $S_1\cup S_2$; intuitivamente, G es la conectiva de disyunción). Sin embargo, es un caso de (consecuencia lógica)$_T$ bajo nuestra selección ordinaria de constantes no lógicas. De forma similar, la oración '$\forall$xx=x', que es una (verdad lógica)$_T$ tanto de *LUn* como de *LAr* bajo la selección ordinaria de constantes lógicas, no es una (verdad lógica)$_T$ bajo la nueva selección.

Como ejemplo de la circunstancia inversa, observemos que la oración '$\forall x \neg Unx$' es (verdad lógica)$_T$ bajo la nueva selección (dado que no hay unicornios y, por tanto, no hay universos U que contengan unicornios), y consiguientemente (consecuencia lógica)$_T$ de todo conjunto de premisas. Sin embargo, no es una (verdad lógica)$_T$ bajo la selección ordinaria. Otro ejemplo: puesto que nadie ha participado nunca como fusilador en su propio fusilamiento (al menos eso supondremos aquí), la oración '$\forall x \neg Fuxx$' será una (verdad lógica)$_T$ bajo la nueva selección y, por tanto, (consecuencia lógica)$_T$ de todo conjunto de premisas; pero no es una (verdad lógica)$_T$ bajo la selección ordinaria.

Pero también podemos observar que estos ejemplos muestran que, bajo la nueva selección extraordinaria, (CLT) y (VLT) son extensionalmente incorrectas. Por un lado, hay casos indiscutibles de consecuencia lógica intuitiva, como el de las inferencias sancionadas por *modus ponens*, que no son casos de (consecuencia lógica)$_T$ bajo la selección extraordinaria de constantes lógicas. Por otro lado, hay casos de (consecuencia lógica)$_T$ bajo la selección extraordinaria que no son casos intuitivos de consecuencia lógica.

Así pues, la cuestión de si las definiciones (CLT) y (VLT) son correctas depende de cuál sea la selección aceptada de constantes lógicas para cada uno de los lenguajes a los que las definiciones son aplicables. Los ejemplos muestran que bajo ciertas selecciones (CLT) y (VLT) son ciertamente incorrectas. Pero no muestran que sean incorrectas bajo toda selección posible. Los argumentos parciales que dimos en favor de su corrección, en los capítulos VII y IX, se basaban implícitamente en la suposición de que un cierto conjunto de constantes, las usuales, que especificamos era el conjunto de las constantes lógicas. (El argumento del capítulo V que zanjaba afirmativamente la primera parte del problema de la corrección no se basaba en una selección de constantes lógicas, sino meramente en la suposición de que dos argumentos tienen la misma forma cuando son obtenibles el uno a partir del otro sustituyendo de manera uniforme constantes no lógicas por constantes no lógicas.) Quizá haya una selección de constantes bajo la cual (CLT) y (VLT) son correctas.

A menudo (aunque no siempre), el concepto de constante lógica se ha introducido como un concepto teórico que surge precisamente de esa hipótesis o, en general, de hipótesis similares en otras teorías "generalistas" de la relación de consecuencia lógica, como la tarskiana. Seamos un poco más explícitos.

Supongamos que tenemos dada una noción de (modo de interpretar)† las constantes de los lenguajes que nos interesan. La hipótesis es, entonces, que hay un conjunto *CoL* de constantes de esos lenguajes tal que un conjunto de premisas K tiene como consecuencia lógica intuitiva a una conclusión X exactamente cuando para todo (modo de interpretar)† las constantes que no están en *CoL* que hace a todas las oraciones de K verdaderas, X es también verdadera.

En el caso de la teoría generalista de Bolzano que mencionamos en el capítulo VII, la versión de esta hipótesis es que hay un conjunto de conceptos (el de los conceptos *lógicos*) tal que un conjunto de premisas K tiene como consecuencia lógica intuitiva a una conclusión X exactamente cuando para toda asignación de conceptos a los conceptos no lógicos que hace a todas las proposiciones de K verdaderas, X es también verdadera. En el caso de la teoría tarskiana, la hipótesis es que hay un conjunto *CoL* de constantes tal que un conjunto de premisas K tiene como consecuencia lógica intuitiva a una conclusión X exactamente cuando para toda estructura apropiada para interpretar las constantes que no están en *CoL* que hace a todas las oraciones de K verdaderas, X es también verdadera. El hipotético conjunto *CoL*, si existe, es el conjunto de las constantes lógicas.

Concentrándonos en la hipótesis correspondiente a la teoría tarskiana, la cuestión de si existe el conjunto *CoL*, y de caracterizarlo de forma iluminadora en el caso de que exista, es una cuestión abierta. De todos modos, tenemos algunas ideas precisas acerca de *CoL*. Una de ellas es que no puede ser vacío; ha de contener al menos constantes como '$\neg$', '$\rightarrow$', '$\forall$', etc. Otra de las ideas que tenemos ya y que debe notarse es que ya contamos con una caracterización, aunque poco iluminadora, que se aplica al conjunto de las constantes lógicas, si existe. A saber, el conjunto de las constantes lógicas ha de verificar la propiedad que por hipótesis se le atribuye: un conjunto de premisas K tiene como consecuencia lógica intuitiva a una conclusión X si y sólo si en toda estructura que interpreta las constantes que no están en ese conjunto y que hace a todas las oraciones de K verdaderas, X es también verdadera.

Estas dos ideas que ya tenemos son suficientes para sacar ciertas conclusiones deseadas. Recordemos nuestra suposición del capítulo IV que explicaba la noción intuitiva de forma en términos de la noción de constante lógica: dos argumentos tienen la misma forma cuando son obtenibles el uno a partir del otro sustituyendo de manera uniforme constantes no lógicas por constantes no lógicas.

Dijimos también (en el capítulo I) que intuitivamente el argumento

El hombre es libre

El hombre tiene responsabilidad moral

y el argumento

El avestruz es comestible

El avestruz tiene sensibilidad estética

parecen tener la misma forma. Y también que el argumento

Doña Pura es soltera

Doña Pura es una mujer no casada

tiene la misma forma que el argumento

Doña Pura es pescadera

Doña Pura es una motocicleta no engrasada

Estas afirmaciones pueden ahora derivarse a partir de las dos ideas acerca de *CoL* enunciadas en el párrafo anterior y la explicación de la noción de (identidad de) forma en términos de la noción de constante lógica.

Lo único que hace falta para establecer esas afirmaciones es mostrar que 'hombre', 'avestruz', 'libre', 'comestible', 'tener responsabilidad moral', 'tener sensibilidad estética', 'soltera', 'mujer', 'casada', 'pescadera', 'motocicleta' y 'engrasada' no son constantes lógicas. (Pues, si no lo son, unos argumentos podrán convertirse en los otros por sustitución uniforme de constantes no lógicas por constantes no lógicas.) Por la primera idea sobre las constantes lógicas, '$\neg$', '$\wedge$', '$\forall$', '$=$', son todas constantes lógicas. Con ayuda de estas constantes, de variables y del predicado 'H', interpretado como constante lógica

y como significando "es un hombre", es posible escribir, para cualquier número natural n, una oración que diga "hay a lo sumo n hombres". Pues bien, si n es un número astronómico, mayor que el número del conjunto de los hombres habidos y por haber, la oración correspondiente será una (verdad lógica)$_T$, una (consecuencia lógica)$_T$ de todo conjunto de oraciones (pues no habrá universos U con más de n hombres como miembros). Pero esa oración no es una consecuencia lógica intuitiva de todo conjunto de oraciones, pues no es implicada por necesidad lógica por todo conjunto de oraciones. Así, por nuestra segunda idea sobre las constantes lógicas, 'H' no pertenece al conjunto de las constantes lógicas. (De forma enteramente análoga puede mostrarse que los otros predicados de la lista no son constantes lógicas. Para mostrar –usando nuestras dos ideas– que ciertos predicados con extensión infinita no son constantes lógicas, usaríamos posiblemente constantes lógicas de lenguajes de orden superior en la construcción de (verdades lógicas)$_T$ que no fueran verdades lógicas intuitivas.)

Las dos ideas sobre las constantes lógicas que hemos descrito y usado no proporcionan una caracterización suficientemente precisa y aclaradora del conjunto de las constantes lógicas. En particular, en la formulación de la segunda de esas ideas usamos el concepto intuitivo de consecuencia lógica, que es el que queremos explicar o iluminar por medio del concepto tarskiano definido de (consecuencia lógica)$_T$. Así pues, sería ocioso emplear esas ideas para explicar la noción de constante lógica que se usa en la descripción general del método tarskiano para definir (consecuencia lógica)$_T$ para lenguajes arbitrarios, descripción general que vimos en el capítulo VI. Tarski, descontento con el hecho de que la descripción general de su método tomara la noción de constante lógica como una noción primitiva, investigó la posibilidad de dar una caracterización o definición precisa del conjunto de las constantes lógicas en términos de conceptos matemáticos. A una discusión del punto de vista de Tarski sobre este problema dedicaremos el próximo capítulo.

XI. La definición de constante lógica de Tarski*

Hay varias ideas explicadas en el capítulo anterior con las que Tarski no habría estado de acuerdo. Como vimos en el capítulo V, Tarski era escéptico acerca del componente modal en la relación de consecuencia lógica intuitiva (y acerca de las nociones modales en general). Por eso habría dudado de la validez de las consideraciones del capítulo anterior que utilizamos para mostrar que expresiones como 'hombre', 'avestruz', 'libre', etc., no son constantes lógicas; pues esas consideraciones hacían un uso esencial de nuestras ideas intuitivas sobre si ciertas oraciones en que aparecían (formalizaciones de) esas expresiones eran implicadas por necesidad lógica por todo conjunto de premisas.

De manera consistente con su escepticismo acerca del componente modal en la relación intuitiva de consecuencia lógica, Tarski también expresó opiniones escépticas acerca de la posibilidad de dar un fundamento filosófico sustancial a la distinción entre constantes lógicas y no lógicas. En "Sobre el concepto de consecuencia lógica", Tarski admitió que la distinción no es enteramente arbitraria, porque si incluimos signos como el cuantificador universal o el condicional entre las constantes no lógicas, "nuestra definición del concepto de consecuencia llevaría a resultados que obviamente contradicen el uso

*Éste capítulo trata temas algo más avanzados que los introductorios; puede omitirse en una primera lectura.

ordinario" (Tarski, 1936: 418). En el capítulo anterior vimos un ejemplo que muestra por qué esto es así; vimos que si considerábamos a '$\rightarrow$' como constante no lógica, entonces argumentos como

$$\frac{(\forall x \neg Unx \rightarrow \exists x \exists x Fuxx) \quad \forall x \neg Unx}{\exists x \exists x Fuxx.}$$

no serían casos de (consecuencia lógica)$_T$, a pesar de que son claramente casos de consecuencia lógica en el sentido intuitivo. Lo que Tarski ponía en duda era que el problema opuesto fuera realmente un problema. Es decir, dudaba que incluir signos como 'hombre', 'avestruz', etc., entre las constantes *lógicas*, haga que la definición de (consecuencia lógica)$_T$ lleve a resultados en obvia contradicción con el uso ordinario de la noción de consecuencia lógica. Según Tarski, incluso si consideramos que todas las constantes son lógicas, la definición ofrece una caracterización de un concepto especial de consecuencia que no es en absoluto desconocido: el de consecuencia material (pues si no hay constantes no lógicas, X es (consecuencia lógica)$_T$ de K cuando alguna de las oraciones de K es falsa o X es verdadera). Para Tarski, hay usos de la expresión 'consecuencia lógica' que abarcan argumentos donde la implicación entre las premisas y la conclusión es "meramente" material.

Cuando Tarski habla de extender de esta manera generosa el conjunto de las constantes lógicas, y de los correspondientes usos de 'consecuencia lógica', parece tener en mente el fenómeno de lo que llama "las disciplinas que preceden a una disciplina dada" (Tarski, 1937: 80). Tarski habla de la lógica como una disciplina que precede a todas las demás, en el sentido de que las constantes lógicas usuales y las leyes lógicas usuales se emplean en todas las disciplinas. Pero, de manera similar, al desarrollar una cierta teoría podemos dar por supuestas no sólo a la lógica sino a otras teorías; así, por ejemplo, Tarski habla de la lógica y la aritmética como una base teórica conveniente para el desarrollo de la geometría. En Gómez Torrente (1996) he indicado que Tarski usaba 'lógica' con distintos grados de amplitud en distintos contextos. Esto se compadecería muy bien con su escepticismo acerca de la posibilidad de encontrar una distinción precisa entre las constantes lógicas y las no lógicas. En el uso de

la expresión característico del propio Tarski, qué constantes y qué leyes se toman como parte de la "lógica" de la investigación depende precisamente del contexto teórico en que se lleva a cabo esa investigación.

Un ejemplo que sabemos que Tarski tenía en mente es el signo para la relación de pertenencia. En una carta de esta época dice:

> "algunas veces me parece conveniente incluir términos matemáticos, como la relación $\in$, en el conjunto de los términos lógicos, y algunas veces prefiero limitarme a los términos de la 'lógica elemental'. ¿Hay algún problema en esto?" (Tarski, 1944: 29).

Según Tarski, podemos tomar el signo de pertenencia como una constante lógica en algunas formalizaciones de la teoría de los tipos (que pueden a su vez ser usadas en la formalización de teorías con universos arbitrarios de individuos). Pero cuando la teoría de conjuntos misma es el objeto de estudio teórico, lo apropiado es no tomar la noción de pertenencia como una noción lógica y no aceptar principios acerca de la noción de pertenencia como principios de la "lógica" de la teoría, sino como postulados de ésta. Las observaciones finales de Tarski en su artículo sobre el concepto de consecuencia lógica expresan de manera clara este punto de vista sobre la relatividad de la noción de constante lógica. Según él, en diferentes contextos se pueden presuponer diferentes conjuntos de constantes lógicas, y puede así variar la extensión del concepto presupuesto de consecuencia lógica. Tarski concluye su artículo diciendo que "la fluctuación en el uso común del concepto de consecuencia quedaría –en parte al menos– reflejada de manera muy natural en esa situación inevitable" (Tarski, 1936: 420).

Sin embargo, Tarski no podía dejar de reconocer que hay ciertos usos privilegiados de la expresión 'consecuencia lógica', y consiguientemente de la expresión 'constante lógica', y, de hecho, este reconocimiento está presupuesto en la enunciación de sus ideas acerca de los diferentes usos de la expresión 'lógica'. Por ello, a Tarski sin duda le pareció insatisfactorio, desde un punto de vista teórico, el haber ofrecido su método para definir (consecuencia lógica)$_T$ para los lenguajes formales y haber dejado sin una definición similar el concepto de constante lógica de esos lenguajes, concepto que aparece como una idea primitiva en la descripción general de su método que vimos en el capítulo

VI. Dejando sin definir el concepto de constante lógica, no se estaría ofreciendo un modo preciso de aislar aquellos argumentos que son casos de consecuencia lógica según el uso privilegiado de 'consecuencia lógica', y esto no parece deseable. Por otro lado, tampoco se estaría ofreciendo un fundamento para incluir signos como el cuantificador universal o el condicional entre las constantes lógicas, y esto es más claramente indeseable. Seguramente por estas razones, Tarski se propuso definir de manera satisfactoria para él el concepto de constante lógica. Una definición de este concepto sería satisfactoria para él si caracterizara un concepto cuya extensión fuera el conjunto de las constantes lógicas, según el uso privilegiado de 'constante lógica', y, además, fuera una definición dada en términos de conceptos científicamente austeros, como los de la lógica y la matemática o la propia semántica tarskiana.

Treinta años después de "Sobre el concepto de consecuencia lógica", en su conferencia "¿Qué son las nociones lógicas?" (1966), Tarski inició un asalto al problema de definir el concepto de constante lógica. En esta conferencia, define el concepto de "noción lógica". En el sentido técnico de Tarski, una "noción" es un objeto de alguno de los tipos posibles en una jerarquía de tipos como la de *Principia Mathematica*: los individuos son nociones, los conjuntos de individuos son nociones, los conjuntos de pares de individuos (relaciones binarias) son nociones, los conjuntos de conjuntos de individuos son nociones, etc. (Véase Tarski, 1966: 147). Tarski propone decir que una *noción lógica* es una noción invariante bajo todas las transformaciones 1-1, o permutaciones, del universo de discurso sobre sí mismo (véase Tarski, 1966: 149).

Una permutación de una clase induce permutaciones de todos los tipos en la jerarquía de tipos de nociones determinada por esa clase. Así, una permutación P de un dominio de individuos D induce una permutación $P_R^{n\,\prime}$ de la clase de las relaciones n-arias de elementos de D; una permutación $P_F^{n\,\prime}$ de la clase de las funciones con n argumentos con dominio D^n y recorrido incluido en D; una permutación $P_R^{n\,\prime\prime}$ de la clase de relaciones n-arias entre relaciones de elementos de D; etc. Una noción u objeto O de un cierto tipo t es invariante bajo todas las permutaciones del universo de discurso si, para todas las permutaciones P de este universo, las permutaciones $\tilde{P}$ inducidas por P en la clase de las nociones de tipo t son todas tales que $\tilde{P}(O)=O$.

En una versión de su conferencia de 1966, Tarski indicó que las funciones veritativas y las denotaciones de los cuantificadores clásicos se pueden construir

en la jerarquía de tipos como ciertos objetos que son invariantes bajo todas las permutaciones del universo de discurso y, por tanto, son nociones lógicas. Por ejemplo, los valores de verdad "verdadero" y "falso" pueden identificarse con el universo de discurso y el conjunto vacío, respectivamente, y las funciones veritativas identificarse a su vez con funciones que tienen (tuplas de) estas clases como argumentos y valores; y las denotaciones de los cuantificadores clásicos universales y existenciales sobre un tipo de objetos t pueden identificarse con ciertas funciones de la clase de los conjuntos de objetos de tipo t en la clase de los valores de verdad –continuando con la identificación de "verdadero" con el conjunto universal de objetos del tipo t y de "falso" con el conjunto vacío de ese tipo. La denotación de un cuantificador universal asignará "verdadero" al conjunto de todos los objetos de tipo t, y "falso" a todos los otros subconjuntos de t; y la denotación del cuantificador existencial asignará "verdadero" a los subconjuntos no vacíos, y "falso" al subconjunto vacío.

En su conferencia de 1966, Tarski no ofrece una definición del concepto de constante lógica. Para encontrar esa definición hay que esperar otros veinte años, a un libro escrito en colaboración con Steven Givant y publicado en 1987, cuatro años después de la muerte de Tarski (Tarski y Givant, 1987). Antes de ver la definición, es conveniente motivarla aproximadamente como sigue. Si queremos definir el concepto de constante lógica en términos del de noción lógica, lo primero que podría ocurrírsenos sería proponer que una constante C es una constante lógica de un lenguaje formal interpretado L si C denota una noción lógica en la interpretación de L. Con esta definición, por ejemplo, y dadas las convenciones sobre la denotación de las conectivas veritativo-funcionales y los cuantificadores que acabamos de ver, estos signos quedan declarados constantes lógicas, no importa cuál sea la interpretación de un lenguaje L que esté en cuestión. Este es un resultado deseado. Sin embargo, esa primera definición de 'constante lógica' sería problemática en otros casos.

Obsérvese que la misma noción aparecerá frecuentemente en las jerarquías de tipos generadas por diferentes dominios, y que en cuanto noción generada por algunos dominios será una noción lógica, mientras que en cuanto noción generada por otros dominios no será lógica. El caso más obvio es el de las clases que se toman como dominios o universos de discurso. Cualquier clase no vacía puede tomarse como universo de discurso, y es una noción

lógica en cuanto noción de la jerarquía de tipos de nociones que la tiene a ella como generadora; pero ninguna clase no vacía es una noción lógica en cuanto noción de la jerarquía generada por una clase de la cual es una subclase propia. Supongamos que hemos adoptado la definición de 'constante lógica' del párrafo anterior. Entonces, si un predicado como el predicado 'N' del lenguaje *LAr* del capítulo IV es verdadero de todos y sólo los individuos en el universo de la interpretación natural de *LAr* (los números naturales), *N* será una constante lógica de *LAr* (y será no lógica para otros modos de interpretar *LAr* cuyos universos incluyan la clase de los números naturales como clase propia). Este resultado es indeseable, incluso para un relativista como Tarski, pues violaría el requisito de que el concepto definido de constante lógica tuviera como extensión el conjunto de constantes determinado por el uso privilegiado de la expresión 'constante lógica'; según éste, 'N' no es una constante lógica de *LAr* interpretado de la manera usual.

Una forma natural de evitar este problema es convertir el concepto relativo de noción lógica, que resulta de la noción de Tarski, en una definición "absoluta" de constante lógica. Una constante lógica de un lenguaje interpretado puede definirse como una constante que denota a una noción lógica en *todo* universo de discurso y, por tanto, en toda interpretación. (Nótese que, para que esta definición, o la tentativa definición anterior, puedan ser mínimamente adecuadas, se está suponiendo que el significado de las constantes está fijo, aunque naturalmente no lo esté su denotación en diferentes interpretaciones. Esto es intuitivamente natural, pues el que una constante sea lógica o no depende de lo que la constante significa. Un signo a secas, independientemente de su significado, no puede ser ni lógico ni no lógico.) Esta definición absoluta es la propuesta por Tarski y Givant.

Seamos un poco más precisos y usemos una notación que nos resultará cómoda posteriormente. Digamos que $den(C, U)$ es la denotación de una constante C en un universo U. (Suponemos, como Tarski y Givant, que el significado de la constante y el universo de discurso bastan para determinar esa denotación.) Entonces la definición de Tarski y Givant puede expresarse así:

(CtLT) Una constante C es una (constante lógica)^T (leído "una constante lógica tarskiana") si y sólo si para todo universo U y toda permutación P de U, $\tilde{P}(den(C, U))=den(C, U)$.

('(CtLT)' significa "constante lógica tarskiana".) Se entiende que el concepto "C denota O en U" puede definirse, para cada lenguaje L, usando los métodos de Tarski (1935). C puede ser una constante individual, una constante de predicado n-aria del segundo tipo, una constante funcional n-aria del segundo tipo, una constante de predicado n-aria del tercer tipo, una constante funcional n-aria del tercer tipo, etc.

La definición (CtLT) está relacionada con una definición basada en una condición más fuerte, que no determina necesariamente la misma extensión para el concepto definido de constante lógica. Dada una biyección B entre dos universos U y V, usemos la notación '$\tilde{B}$' para designar la biyección inducida por B entre los objetos de la jerarquía de tipos generada por U y los objetos de la jerarquía de tipos generada por V. Entonces, la definición basada en la condición más fuerte es:

(CtLM) Una constante C es una (constante lógica)^M (leído "una constante lógica Mostowskiana") si y sólo si para cualesquiera universos U y V de la misma cardinalidad y toda biyección B de U sobre V, $\tilde{B}(den(C, U))=den(C, V)$.

('(CtLM)' significa "constante lógica Mostowskiana". Uso esta terminología ya que Mostowski (1957) dio una definición de una cierta clase de cuantificadores generalizados de primer orden como, en esencia, los cuantificadores cuya denotación es una (constante lógica)^M. Pero hay que subrayar que Mostowski no usó su condición de invariancia bajo biyecciones para dar una definición del concepto de constante lógica. Esta condición ha sido usada para este propósito por otros autores, por ejemplo Sher (1991).) Dicho de manera intuitiva, (CtLM) requiere que las constantes lógicas denoten, en todo universo, no meramente una noción invariante bajo permutaciones de ese universo, sino una noción que sea la misma, salvo isomorfismo, en *todos* los universos de la misma cardinalidad. Si una constante es una (constante lógica)^M entonces es una (constante lógica)^T, pues toda permutación de un universo U es una biyección de U sobre U. Pero no es necesariamente el caso que toda (constante lógica)^T sea una (constante lógica)^M. Supongamos, por ejemplo, que *hubiera* una constante de predicado monádica del segundo tipo, C, cuyo significado fuera tal que, en todo universo, C denotara o bien el conjunto vacío

o el universo mismo, y que denotara el conjunto vacío en algunos universos de una cierta cardinalidad κ y el conjunto universal en otros universos de cardinalidad κ. *C* sería una (constante lógica)^{T}, pero no una (constante lógica)^{M}.

Tanto (CtLT) como (CtLM) satisfacen el requisito, deseado por Tarski, de que una definición del concepto de constante lógica habría de darse en términos de conceptos austeros y ciertamente no modales, como los de la lógica, la matemática y la semántica tarskiana. De esta manera, al eliminar el concepto preteórico de constante lógica de la descripción del método tarskiano para definir (consecuencia lógica)$_{T}$ y sustituirlo por uno de los conceptos definidos en (CtLT) o (CtLM), el resultado sigue siendo un concepto definido de (consecuencia lógica)$_{T}$ que explica la noción preteórica de consecuencia lógica en términos de un aparato de conceptos mejor comprendidos que esa noción preteórica o que la noción de implicación por necesidad lógica. La cuestión de si (CtLM) y (CtLM) satisfacen el otro requisito de Tarski, el de determinar la extensión determinada por el uso privilegiado de 'constante lógica', es una cuestión más espinosa. A ella dedicaremos la primera parte del próximo capítulo.

XII. UN EXAMEN DEL CONCEPTO TARSKIANO DE CONSTANTE LÓGICA*

Tarski y Lindenbaum demostraron en 1935 que, dado un universo básico *U*, todas las nociones de la jerarquía de nociones generada por *U* que pueden ser definidas en el lenguaje de la teoría simple de los tipos son invariantes bajo todas las permutaciones de *U*. Es decir, las relaciones de individuos, relaciones de relaciones de individuos, etc., que pueden definirse en la teoría de los tipos no cambian tras ninguna permutación del universo de individuos. Por ejemplo, ningún individuo es una noción lógica, pues todo individuo puede proyectarse sobre uno diferente en una permutación (si hay más de un individuo en el universo); y ningún individuo es definible en la teoría de tipos. Las clases de individuos definibles en la teoría de los tipos son la clase de todos los individuos y la clase vacía, y éstas son las únicas clases de individuos invariantes bajo todas las permutaciones del universo. Y, así, sucesivamente.

En general, el teorema de Tarski y Lindenbaum garantiza que todas las nociones matemáticas definibles a la manera logicista en la teoría simple de los tipos son nociones lógicas, no importa cuál sea el universo supuesto de individuos. Puesto que el teorema vale para todo universo *U* que proporcione una interpretación del lenguaje de la teoría de los tipos, la definición (CtLT)

*Este capítulo trata temas algo más avanzados que los introductorios; puede omitirse en una primera lectura.

implica que todos los símbolos primitivos que denoten nociones en ese lenguaje (por ejemplo, cuantificadores de todos los tipos) son (constantes lógicas)^{T}; además, si la definición se aplicara a símbolos definidos, todos estos símbolos serían (constantes lógicas)^{T}. Estos resultados se compadecen bien con el uso de la expresión 'constante lógica' por los logicistas, y por otros lógicos y filósofos (si bien no por todos, pues, como ya dijimos, muchos no aceptan que los cuantificadores de órdenes superiores al primero sean constantes lógicas).

Parece razonable pensar que hay, en esencia, dos usos típicos de la expresión 'constante lógica': el uso característico de los logicistas y un uso, menos generoso, que excluye a los cuantificadores de órdenes superiores al primero del conjunto de las constantes lógicas. Si algún uso de 'constante lógica' es el privilegiado, entonces es uno de estos dos. En general, el teorema de Tarski y Lindenbaum nos da una buena justificación para pensar que, sea cual sea de estos dos el uso privilegiado de 'constante lógica', la definición (CtLT) proporciona una condición *necesaria* de las constantes lógicas usuales: todas las constantes lógicas según el uso logicista, el más generoso, son (constantes lógicas)^{T}. Por tanto, todas las constantes lógicas en cualquier uso menos generoso han de ser (constantes lógicas)^{T}.

La cuestión inversa, la de si la definición (CtLT) proporciona una condición *suficiente* para que una (constante lógica)^{T} pertenezca al conjunto usual de constantes lógicas, parece que ha de tener una respuesta negativa, independientemente de cuál sea el uso de la expresión 'constante lógica' que adoptemos como privilegiado. Cuando (CtLT) se aplica a lenguajes interpretados de teorías con primitivos que habitualmente se toman como no lógicos (por ejemplo, lenguajes de primer orden como *LAr* o como *LUn*), el resultado generalmente será que las nociones denotadas por estos primitivos no son (constantes lógicas)^{T}. Así, los símbolos '0', 'N', 'M' de *LAr*, los símbolos 'π' y 'Fu' de *LUn* y símbolos formales para 'hombre', 'avestruz', etc. (todos ellos con sus significados deseados, naturalmente) no son (constantes lógicas)^{T}. La razón de esto es que para cada uno de esos símbolos hay universos en que el símbolo en cuestión denota una noción no lógica, como puede comprobarse en un sencillo ejercicio. Estos resultados de nuevo se compadecen bien con el uso habitual de la expresión 'constante lógica' (logicista o no), y confirman parcialmente la hipótesis de que la condición (CtLT) es suficiente para la pertenencia en el conjunto usual de constantes lógicas.

Sin embargo, consideremos el predicado 'Un' del lenguaje *LUn*. Este predicado denota una noción lógica en todos los universos, a saber, el conjunto vacío (pues no hay unicornios). Por tanto, es una (constante lógica)^{T}. (Y también una (constante lógica)^{M}.) Pero en ninguno de los usos habituales de 'constante lógica' se aceptaría a 'Un' como constante lógica. Naturalmente, hay muchos otros ejemplos similares. Estos ejemplos parecen mostrar concluyentemente que la definición de constante lógica de Tarski (y también la inspirada en las ideas de Mostowski) no cumple con el requisito de determinar el mismo conjunto de constantes que el uso privilegiado de la expresión 'constante lógica'.

Es conveniente mencionar que una cierta modificación de (CtLT) (y de (CtLM)), que podría ocurrírsenos inmediatamente, tampoco cumple con los requisitos tarskianos para una definición del concepto de constante lógica (con ninguno de los dos). La modificación consistiría en requerir, en el caso de (CtLT), que una (constante lógica)^{T} denote nociones lógicas en todos los universos *posibles* (en el caso de (CtLM), consistiría en requerir que una (constante lógica)^{M} denote objetos invariantes salvo isomorfismo en universos *posibles* de la misma cardinalidad). Las nuevas definiciones resultantes (de (constante lógica)$^{T^*}$ y (constante lógica)$^{M^*}$, digamos) no cumplirían con el requisito de darse en términos aceptables, desde un punto de vista científico y escéptico con las nociones modales, pues utilizan el concepto de universo *posible*. Pero ni siquiera cumplirían con el otro requisito tarskiano. Incluso si suponemos que hay universos posibles en los que el conjunto de unicornios no es vacío (algo que, como es sabido, niegan algunos filósofos del lenguaje contemporáneos), es posible hallar símbolos que denoten *necesariamente* nociones lógicas (y objetos invariantes salvo isomorfismo en todos los universos *posibles* de la misma cardinalidad). Definamos, por ejemplo, un *heptaedro* como un poliedro regular de siete caras. No hay heptaedros en ningún mundo posible, de manera que un predicado formal 'Hp' con el mismo significado que 'heptaedro' sería una (constante lógica)$^{T^*}$ (y una (constante lógica)$^{M^*}$), a pesar de que en ningún uso habitual de la expresión 'constante lógica' se consideraría a 'Hp' como una constante lógica.

Así pues, la definición (CtLT) (y también (CtLM)) parece enfrentarse a obstáculos muy considerables. No hay manera obvia de corregirla de modo que cumpla con los requisitos tarskianos para una definición del concepto de

consecuencia lógica. De todos modos, es importante observar que una definición del tipo de la propuesta por Tarski no es *necesariamente* problemática en otro sentido en que se pudiera pensar que lo es. Como hemos venido subrayando repetidamente desde el capítulo V, el razonamiento intuitivo mixto con nociones matemáticas y nociones modales es posible y a menudo puede usarse para establecer conexiones importantes entre nociones como las definidas por Tarski y nociones modales intuitivas (aunque, como sabemos, Tarski habría visto con escepticismo este tipo de razonamiento). Hemos visto numerosos ejemplos de este tipo de razonamiento en esta monografía. Pues bien, es importante observar que nada impide el uso de este tipo de razonamiento para establecer conexiones interesantes entre un concepto de constante lógica como el definido matemáticamente por Tarski y nociones modales intuitivas. Al establecimiento de una de estas conexiones dedicaremos lo que resta de este capítulo.

Como vimos, la parte crucial del problema de la corrección de la definición tarskiana de consecuencia lógica consistía en ofrecer razones para pensar que los argumentos que son casos de (consecuencia lógica)$_T$ son también casos de consecuencia lógica intuitiva. En particular, para pensar que son casos de implicación por necesidad lógica, noción ésta que se podía elucidar por medio de varias relaciones modales que parecen dar, cuando menos, propiedades necesarias de los casos de implicación por necesidad lógica, como la de implicación analítica, la de implicación *a priori*, la de (consecuencia lógica)$_{Cl}$, etc. Cuando investigamos el problema de la corrección, lo hicimos siempre bajo el supuesto de que las constantes lógicas son las habituales dado un uso habitual desde los logicistas. Pero una vez que cuenta uno con la definición de (constante lógica)^T (o de (constante lógica)^M), es natural plantearse el problema de la corrección de la definición tarskiana de (consecuencia lógica)$_T$ cuando por 'constante lógica' uno entiende (constante lógica)^T (o (constante lógica)^M).

Como era de esperar por nuestras consideraciones de más arriba, la definición tarskiana de (consecuencia lógica)$_T$ parece claramente incorrecta cuando por 'constante lógica' uno entiende (constante lógica)^T (o (constante lógica)^M). Oraciones como '$\forall x \neg Unx$', del lenguaje LUn, son (consecuencia lógica)$_T$ de todo conjunto de oraciones (y por tanto (verdades lógicas)$_T$) cuando por 'constante lógica' uno entiende (constante lógica)^T (o (constante lógica)^M), situación en la cual 'Un' queda declarada constante lógica. Sin embargo, esas oraciones claramente no son implicadas analíticamente ni *a priori* por todo

conjunto de oraciones. Y tampoco son (consecuencias lógicas)$_{Cl}$ o (consecuencias lógicas)$_{CoP}$ de todo conjunto de oraciones (pues, por ejemplo, hay conjuntos de unicornios posibles; o al menos eso se puede suponer).

Sin embargo, es posible establecer que la definición tarskiana de (consecuencia lógica)$_T$ es parcialmente correcta cuando por 'constante lógica' uno entiende (constante lógica)^M, en el sentido de que es posible establecer que ciertos casos de (consecuencia lógica)$_T$ (cuando por 'constante lógica' uno entiende (constante lógica)^M) que pertenecen a lenguajes con determinadas propiedades son casos de (consecuencia lógica)$_{CoP}$. El valor de este resultado es quizá escaso, en vista de los problemas de las nociones tarskiana y mostowskiana de constante lógica. Pero tiene, cuando menos, un valor metodológico: con él se ilustra que el razonamiento mixto con nociones matemáticas y modales es posible, siendo, en este caso, la noción matemática una noción definida de constante lógica.

El resultado es el siguiente. Supongamos que tenemos un lenguaje L todas cuyas (constantes lógicas)^M son (constantes lógicas)^M *en todos los mundos posibles*. (Las expresiones aceptadas como constantes lógicas en los lenguajes de órdenes primero y superior en el uso logicista tienen esta propiedad; también la tiene la constante 'Hp' que hemos visto hace unos párrafos.) Entonces es posible demostrar que para todo conjunto de oraciones K de L y toda oración X de L, si X es (consecuencia lógica)$_T$ de K (cuando por 'constante lógica' uno entiende (constante lógica)^M), entonces X es (consecuencia lógica)$_{CoP}$ de K.

El hecho fundamental usado en esta demostración es

> *(1) Para toda (estructura)^M, hay una (estructura)^M (isomorfa)^M cuyo dominio es un conjunto puro.*

(Una (estructura)^M para un lenguaje L es una estructura que interpreta las constantes que no son (constantes lógicas)^M de L. La noción de isomorfismo entre (estructuras)^M, para la que uso la notación '(isomorfa)^M', se define de la manera obvia.) La prueba de (1) consiste en observar que si E es una (estructura)^M, entonces su universo tiene un buen orden, y es, por tanto, biyectable con un segmento inicial α de los ordinales (que construimos de nuevo como conjuntos puros); esa biyección induce un isomorfismo entre E y una (estructura)^M con universo α.

Un segundo hecho que necesitamos es

(2) Para cualesquiera dos (estructuras)^{M} (isomorfas)^{M} E y E' y para toda oración O, O es verdadera en E si y sólo si O es verdadera en E'.

La prueba de (2) procedería por inducción sobre la complejidad de las oraciones de cada lenguaje particular, *L* en nuestro caso. La idea intuitiva es clara: si *E* y *E'* son (isomorfas)^{M} entonces (CtLM) garantiza que son "imágenes reflejas" no sólo con respecto a sus universos y a las denotaciones de las constantes de *O* que no son (constantes lógicas)^{M}, sino también con respecto a las denotaciones de las constantes de *O* que son (constantes lógicas)^{M}.

El razonamiento es entonces como sigue. Supongamos que *X* es (consecuencia lógica)$_{T}$ de *K* (cuando por 'constante lógica' uno entiende (constante lógica)^{M}), pero *X* es falsa en un (modo de interpretar)$_{CoP}$ *I*, existente en un mundo posible *m*, donde todas las oraciones de *K* son verdaderas. (1) es verdadero en todo mundo posible y, por tanto, en *m*. Luego hay un (modo de interpretar)$_{CoP}$ *I'*, existente en *m*, con un conjunto puro como universo, que es (isomorfa)^{M} a *I*. Pero también (2) es verdadero en *m*. Puesto que estamos suponiendo que las constantes de *L* que son (constantes lógicas)^{M} (en el mundo actual) son (constantes lógicas)^{M} en *m*, se sigue de (2) que *X* ha de ser falsa y las oraciones de *K* verdaderas en *I'*. Dado que *I'* existe también en el mundo actual, *X* no es (consecuencia lógica)$_{T}$ de *K* (cuando por 'constante lógica' uno entiende (constante lógica)^{M}). Y esto contradice nuestro supuesto inicial.

APÉNDICE. LAS SEMICOMILLAS

Al dar una definición tarskiana del predicado 'es (consecuencia lógica)$_{T}$ de' para un lenguaje, es necesario no sólo poder referirnos mediante nombres por entrecomillado a expresiones particulares de ese lenguaje, sino también poder hacer generalizaciones acerca de *todas* (o *algunas*) de las expresiones del lenguaje que tienen una cierta forma en común.

Por ejemplo, necesitamos poder decir no sólo que la expresión particular '$\forall x_n x_n = x_n$' de *LAr* tiene tal o cual propiedad (verbigracia, la propiedad de ser una oración), sino también cosas tales como que *para toda* variable x_n, la expresión que resulta de concatenar '$\forall$' con x_n con x_n con '=' con x_n, en ese orden, tiene tal propiedad (verbigracia, la propiedad de ser una oración de *LAr*).

Observemos que para decir que para toda variable x_n, la expresión que resulta de concatenar '$\forall$' con x_n con x_n con '=' con x_n, en ese orden, es una oración de *LAr*, no podemos usar un nombre por entrecomillado; no podemos decir: para toda variable x_n, '$\forall x_n x_n = x_n$' es una oración de *LAr*. La razón es que

'$\forall x_n x_n = x_n$'

es un nombre de una expresión *particular*, que ni siquiera es una oración de *LAr*. Por tanto, el efecto de generalizar sobre todas las expresiones de una cierta forma no se consigue.

Otro ejemplo. Cuando en lógica decimos informalmente que *toda* fórmula proposicional de la forma A→A es una verdad lógica, lo que estamos diciendo no puede ponerse diciendo: para toda fórmula proposicional A, 'A→A' es una verdad lógica. De esta forma no conseguimos generalizar sobre toda fórmula proposicional; simplemente estamos haciendo una cuantificación vacua: para toda fórmula proposicional A, la expresión *particular* 'A→A' es una verdad lógica (y 'A→A' ni siquiera es una oración de un lenguaje proposicional, pues 'A' no es una fórmula proposicional, sino una variable de nuestro metalenguaje para hablar de la lógica proposicional). Para conseguir el efecto deseado debemos decir, por ejemplo: para toda fórmula proposicional A, el resultado de concatenar A con '→' con A, en ese orden, es una verdad lógica.

Al expresar con estricta corrección lo que queremos decir en estos casos, es muy conveniente tener formas perspicuas de abreviar expresiones demasiado largas como 'la expresión que resulta de concatenar '∀' con x_n con x_n con '=' con x_n, en ese orden'. Un método para abreviar tales expresiones que es especialmente perspicuo se debe a Quine. La idea de Quine consiste simplemente en usar, donde nos vemos tentados a usar comillas, un par de símbolos tipográficamente parecidos a las comillas ('⌜', '⌝'), pero con el significado deseado; a veces se llama a estos símbolos 'semicomillas'.

Usemos nuestros ejemplos para ver cómo funcionan las semicomillas. En el primer caso, abreviamos la expresión 'la expresión que resulta de concatenar '∀' con x_n con x_n con '=' con x_n, en ese orden' por medio de la expresión $\ulcorner\forall x_n x_n = x_n\urcorner$. Podemos entonces generalizar y decir que para toda variable x_n, $\ulcorner\forall x_n x_n = x_n\urcorner$ es una oración de *LAr*. En el segundo caso, abreviamos 'el resultado de concatenar A con '→' con A, en ese orden' por medio de $\ulcorner A \rightarrow A\urcorner$. Hacemos entonces nuestra generalización y decimos: para toda fórmula proposicional A, $\ulcorner A \rightarrow A\urcorner$ es una verdad lógica.

REFERENCIAS BIBLIOGRÁFICAS

Boolos, G. (1975): "On Second-Order Logic", *Journal of Philosophy*, vol. 72.

— (1984): "To Be is to Be a Value of a Variable (or to Be some Values of some Variables)", *Journal of Philosophy*, vol. 81.

Etchemendy, J. (1990): *The Concept of Logical Consequence*, Harvard University Press, Cambridge (Massachusetts).

Gómez Torrente, M. (1996): "Tarski on Logical Consequence", *Notre Dame Journal of Formal Logic*, vol. 37.

— (1996a): *Tarski's Definition of Logical Consequence. Historical and Philosophical Aspects* (tesis doctoral, Princeton University), University Microfilms, Ann Arbor (Michigan).

— (1998): "On a Fallacy Attributed to Tarski", *History and Philosophy of Logic*, vol. 19.

— (1998/9): "Logical Truth and Tarskian Logical Truth", *Synthese*, vol. 117.

Kreisel, G. (1967): "Informal Rigour and Completeness Proofs", en I. Lakatos (ed.), *Problems in the Philosophy of Mathematics*, North-Holland, Amsterdam.

Maddy, P. (1988): "Believing the Axioms", *Journal of Symbolic Logic*, vol. 53.

Mates, B. (1965): *Lógica matemática elemental*, Tecnos, Madrid, 1970. (Versión inglesa original: *Elementary Logic*, Oxford University Press, Oxford, 1965.)

McGee, V. (1992): "Two Problems with Tarski's Theory of Consequence", *Proceedings of the Aristotelian Society*, n.s., vol. 92.

Mostowski, A. (1957): "On a Generalization of Quantifiers", *Fundamenta Mathematicae*, vol. 44.

Shapiro, S. (1991): *Foundations without Foundationalism: a Case for Second-Order Logic*, Clarendon Press, Oxford.

Sher, G. (1991): *The Bounds of Logic. A Generalized Viewpoint*, M. I. T. Press, Cambridge (Massachusetts).

Tarski, A. (1935): "The Concept of Truth in Formalized Languages", en Tarski (1983).

— (1936): "On the Concept of Logical Consequence", en Tarski (1983).

— (1937): *Einführung in die mathematische Logik und in die Methodologie der Mathematik*, Springer, Viena.

— (1944): "A Philosophical Letter of Alfred Tarski", *Journal of Philosophy*, vol. 84, 1987.

— (1966): "What Are Logical Notions?", *History and Philosophy of Logic*, vol. 7, 1986.

— (1983): *Logic, Semantics, Metamathematics*, 2ª edición, Hackett, Indianapolis.

Tarski, A. y S. Givant (1987): *A Formalization of Set Theory without Variables*, American Mathematical Society, Providence (Rhode Island).

Wang, H. (1974): "The Concept of Set", en P. Benacerraf y H. Putnam (eds.), *The Philosophy of Mathematics*, 2ª edición, Cambridge University Press, Cambridge, 1983.

OTROS TÍTULOS DE LA COLECCIÓN

Lógica informal
Juan Manuel Comesaña

La verdad desestructurada
Eduardo Alejandro Barrio

Concepciones de la referencia
Eleonora Orlando

Conjuntos e infinitos
Ana Carolina Sartorio